DOCUMENT RESUME

ED 147 016 PS 009 658

AUTHOR	Travers, Jeffrey; And Others
TITLE	National Day Care Study Second Annual Report, 1975-1976: Phase II Results and Phase III Design.
INSTITUTION	Abt Associates, Inc. Cambridge, Mass.
SPONS AGENCY	Office of Child Development (DHEW), Washington, D.C. Day Care Services Div.; Office of Human Development (DHEW), Washington, D.C.
PUB DATE	Apr 77
CONTRACT	HEW-105-74-1100
NOTE	111p.; For First Annual Report, see ED 131 928-930
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	*Annual Reports; Class Size; *Day Care Programs; *Early Childhood Education; *Educational Policy; *Educational Research; *Federal Programs; Research Design; Research Methodology; Student Teacher Ratio; Teacher Qualifications
IDENTIFIERS	National Day Care Study; *Staff Child Ratio

ABSTRACT

This second annual report of the National Day Care Study (NDCS), a 3-year study funded by the Office of Child Development, describes the work of the NDCS during Phase II of the study (Sept. 1975 through Sept. 1976). The study was designed to answer policy questions concerning the effect of child ratio and other regulatable center characteristics on child development, per child cost, and cost effectiveness of centers. Chapter 1 of this report discusses the design and sample of the study. Policy variables are explored in Chapter 2. They include staff qualifications, staff/child ratio and group size, and other regulatable center characteristics (physical environment, program orientation, supplemental services, etc.). Chapter 3 discusses the selection of child variables to be measured and instruments used, observation of caregivers and children, and parent interviews. Chapter 4 presents a status report of the financial analysis of the study. Chapter 5 presents Phase III design and data collection procedures. Included in Phase III will be a 49-center quasi-experiment and an eight-center experiment. Appendices include NDCS measurement and analysis methodology, descriptions of the national supply study and infant care study, and lists of the consultant panel, OCD staff and consultants and NDCS staff. (SB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED147016

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

National Day Care Study

SECOND ANNUAL REPORT
1975-1976

Phase II Results and Phase III Design

Principal Authors:	*With:*
Jeffrey Travers	William Bache III
Craig Coelen	David Connell
Richard Ruopp	Frederic Glantz
	Nancy Goodrich
	Robert Goodrich
	Barbara Goodson
	Kathryn Hewett
	Jean Layzer

Prepared for:
Day Care Services Division
Office of Child Development
Office of Human Development
Department of Health, Education,
and Welfare
Washington, D.C.

Government Project Officer:
Allen N. Smith

Contract No.: HEW-105-74-1100

By:
Abt Associates Inc.
Cambridge, Mass. 02138

Study Project Director:
Richard R. Ruopp

PS 009658

Abt Associates, 55 Wheeler Street, Cambridge, Massachusetts 617 492-7100

This report has been reviewed for general publication and meets the quality requirements set by the Research Design Group of Abt Associates. The statements and conclusions contained herein are those of Abt Associates and do not necessarily reflect the views of the sponsoring agency.

Richard R. Ruggs Project Director

Technical Reviewer

Management Reviewer

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS . . . i

INTRODUCTION . . . iii
  BACKGROUND AND GOALS OF THE NATIONAL DAY CARE STUDY . . . iii
  OVERVIEW OF THE STUDY . . . iv
  ORGANIZATION OF THE SECOND ANNUAL REPORT . . . vi
  PURPOSES AND LIMITATIONS OF THE SECOND ANNUAL REPORT . . . vii

Chapter 1: THE PHASE II DESIGN AND SAMPLE . . . 1
  PHASE II DESIGN . . . 1
  DESCRIPTION OF THE PHASE II SAMPLE . . . 4
  EVALUATION OF THE SAMPLE . . . 7
    Variation in Staff/Child Ratio, Group Size and Staff Qualifications . . . 7
    Relations Among the Policy Variables . . . 9
    Relationships Between the Policy Variables and Other Center Characteristics . . . 10
    Implications for Phase II Analysis and Phase III Design . . . 10

Chapter 2: FIDCR POLICY VARIABLES . . . 13
  INTRODUCTION . . . 13
    The Goals and Rationale of the FIDCR . . . 13
    Issues of Monitoring and Measurement . . . 15
  STAFF QUALIFICATIONS . . . 16
    Education . . . 17
    Experience . . . 17
    Job-Related Training . . . 17
    Comparison of Lead Caregivers, Aides and Directors . . . 20
    Implications . . . 21
  STAFF/CHILD RATIO AND GROUP SIZE . . . 21
    Who Should Be Included in Ratio and Group Size Computations? . . . 23
    Should Ratio and Group Size Be Assessed by Records or Observations? . . . 23
    Comparison of Observed and Scheduled Staff/Child Ratios . . . 24
    Should Measures of Ratio and Group Size Be Prospective or Retrospective? . . . 26
    When Should Ratio Be Measured? . . . 26
    How Should Average Ratios Be Computed? . . . 28
  OTHER REGULATABLE CENTER CHARACTERISTICS . . . 30
    The Physical Environment . . . 30
    Center Philosophy and Program Orientation . . . 31
    Supplemental Services and Specialists . . . 32
    Opportunities for Parent Involvement . . . 33
    Continuity of Care . . . 34
    Summary . . . 34

Chapter 3: PROCESSES AND OUTCOMES IN DAY CARE . . . 35
  SELECTION OF CHILD VARIABLES AND INSTRUMENTS . . . 36
  OBSERVATIONS OF CAREGIVERS AND THE CLASSROOM ENVIRONMENT . . . 41
    Physical Environment Inventory . . . 41
    Classroom Snapshot . . . 42
    Five-Minute Interactions . . . 42

OBSERVATIONS OF CHILDREN ... 45
TESTS AND RATINGS OF SOCIOEMOTIONAL AND COGNITIVE DEVELOPMENT ... 49
  The Preschool Inventory (PSI) ... 49
  McCarthy Scale of Verbal Memory (MVM) ... 50
  Matching Familiar Figures Test (MFFT) ... 50
  Motor Inhibition Test (MIT) ... 50
  Pupil Observation Checklist (POCL) ... 50
  Caregiver Ratings ... 51
  Summary ... 52
PARENT INTERVIEWS ... 52
  Family Use of Day Care ... 52
  Parental Expectations ... 53
  Parent Involvement in Center Activities ... 54
  Parent Satisfaction ... 55
  Perceived Impact of the Center on Child and Family ... 55
  Implications ... 56

Chapter 4: NDCS FINANCIAL ANALYSIS ... 59
RESEARCH COST ACCOUNTING SYSTEM (RCAS) ... 59
SUMMARY OF FINANCIAL CHARACTERISTICS OF NDCS CENTERS ... 60
THE NDCS FINANCIAL MODEL ... 62
RELATIONSHIP OF THE FINANCIAL ANALYSIS TO OTHER COMPONENTS OF THE NDCS ... 66

Chapter 5: PHASE III DESIGN AND DATA COLLECTION ... 69
THE 49-CENTER QUASI-EXPERIMENT ... 69
  Sites and Centers ... 70
  Logic of the Design ... 71
  Use of Entering Child Test Scores as an Independent Variable ... 71
  Implementation of the Design ... 72
THE EIGHT CENTER EXPERIMENT ... 72
  The Atlanta Public School Centers ... 73
  Implementation of the Design ... 74
PHASE III MEASUREMENT AND DATA COLLECTION ... 75
  Phase III Child-Focus Observations ... 76
  Phase III Adult-Focus Observations ... 76
  Center Organization, Operations and Cost Data ... 77
  Periodic Center Information ... 77
  Parent Interview and Survey ... 77
  Day Care Forces Inventory ... 77
  Site Case Studies ... 78

Appendix A: NDCS MEASUREMENT AND ANALYSIS METHODOLOGY ... 79

Appendix B: NATIONAL SUPPLY STUDY ... 87

Appendix C: INFANT CARE STUDY ... 91

Appendix D: NDCS CONSULTANT PANEL, OFFICE OF CHILD DEVELOPMENT STAFF AND CONSULTANTS, AND NDCS STAFF ... 93

NOTES TO THE CHAPTERS ... 97

ACKNOWLEDGEMENTS

The National Day Care Study (NDCS) is being conducted over a three-year period by the Day Care Services Division, Office of Child Development (OCD) of the Office of Human Development (OHD), U.S. Department of Health, Education and Welfare (HEW). This *Second Annual Report* is designed to acquaint the reader with the work of the NDCS during its Phase II Natural Study Year, September 1975 through September 1976.

This public report is based largely on the *NDCS Phase II Research Report,* which was the work of the following study staff: William Bache, Craig Coelen, David Connell, Frederic Glantz, Nancy Goodrich, Robert Goodrich, Barbara Goodson, Janet Grasso, Kathryn Hewett, Jean Layzer, Janet Smith, Jan Swanson, Barbara Tannen and Donna Warner.

Jeffrey Travers, NDCS Associate Project Director, drafted most of the *Second Annual Report.* He was assisted substantially with Chapter 1 by Craig Coelen, with Chapter 2 by Richard Ruopp and Craig Coelen, and with Chapter 4 by Frederic Glantz. Anthony Bryk and Craig Coelen drafted Appendix A based on original work by Robert Goodrich, Director of Research. Editorial assistance was provided by Craig Coelen, William C. Hamilton, Christine Jerome, Arthur Mathis, Richard Mayer, Richard Ruopp, and Sally Weiss.

The report was systematically reviewed by three members of the NDCS advisory group of consultants: Asa Hilliard, John Love and James Young.* Allen Smith, the OCD government project officer, critiqued each working draft and made substantive contributions. He was assisted in his efforts by Carol Spence, a consultant to OCD for the NDCS.

Richard R. Ruopp, Project Director
Abt Associates Inc.
Cambridge, Massachusetts
April 1977

**Background information on members of the NDCS consultant panel, OCD staff and consultants, and project staff will be found in Appendix D.*

INTRODUCTION

BACKGROUND AND GOALS OF THE NATIONAL DAY CARE STUDY

In recent years the federal government has become the major funding source for the purchase of day care services for children from six months to fourteen years of age. During the 1976 fiscal year, for example, the Department of Health, Education, and Welfare (HEW) reported spending $743 million under Titles IV-A, IV-B and XX of the Social Security Act to purchase welfare- and work-related care for almost 1.3 million children. Additional HEW funds were spent on day care in connection with Head Start and with various programs funded under the Elementary Education Act.[1]*

As a result of its role as a purchaser of day care, the federal government is concerned both with the costs of care and the effects of different care arrangements on children and parents. In 1968, HEW, the Department of Labor and the Office of Economic Opportunity adopted a uniform set of standards intended to control the type and quality of care purchased with federal dollars. These standards, the Federal Interagency Day Care Requirements (FIDCR), apply to all day care centers and family day care homes serving federally subsidized children. Designed to protect children from harm and to promote their development, the 1968 FIDCR cover a variety of center and family day home characteristics of which three are particularly important to the National Day Care Study: staff/child ratio, group size and staff qualifications.**

The 1968 FIDCR were made part of the Code of Federal Regulations in early 1969. Late in 1974, Congress enacted an amendment to the Social Security Act – Title XX, Grants to States for Services – which slightly modified the 1968 version of the FIDCR and made it law, effective October 1, 1975, with potentially severe financial penalties imposed for noncompliance. The modification added staff/child ratio requirements for children under three and made optional the inclusion of an identifiable educational component. Because of controversy at the state and local level, Congress delayed full implementation of the FIDCR staff/child ratio requirements; centers were not required to meet the federally prescribed ratios for children aged six weeks to six years. At the same time, states were prohibited from allowing ratios in subsidized centers to fall below actual 1975 levels. As of this writing, enforcement of the ratio requirements is not scheduled to begin until the summer of 1977 and is likely to be delayed further.

The 1968 FIDCR were based on expert judgment and the best evidence available at the time. However, in 1968 there existed no carefully designed, large-scale studies illuminating the relationships among FIDCR-regulated center characteristics, the costs of day care, processes in the day care classroom and outcomes for children and their families. The main purpose of the National Day Care Study (NDCS) is to determine the costs and effects associated with different levels of staff/child ratio, group size and staff qualifications. These center characteristics are termed "policy variables" throughout this report. The study is concerned primarily with publicly subsidized preschoolers in larger full-day, full-year urban day care programs.

The NDCS is designed to answer the following major policy questions regarding the three policy variables:

- How is the development of preschool children in federally subsidized day care centers affected by variations in staff/

Numbers in the text refer to notes appearing at the end of this volume. Notes contain full references for all cited sources and statistics and, where appropriate, additional data or discussion.

***In the NDCS* **First Annual Report**[2] *the term "professionalism" was used to label the third major category of regulated center characteristics. Early Phase II results indicated that "qualifications," defined as education, employment experience and job-related training, was more precise from both a research and regulatory perspective.*

child ratio, staff qualifications, group size and/or other regulatable center characteristics?

- How is the per-child cost of center-based day care affected by variations in staff/child ratio, staff qualifications, group size and other regulatable center characteristics?
- How does the cost-effectiveness of federally subsidized, center-based day care change when adjustments are made in staff/child ratio, staff qualifications, group size and other regulatable center characteristics?

OVERVIEW OF THE STUDY

HEW's Office of Child Development (OCD) has funded two research organizations to assist the OCD Day Care Division in conducting the NDCS: Abt Associates Inc. (AAI) of Cambridge, Massachusetts, and Stanford Research Institute (SRI) of Menlo Park, California. AAI has overall administrative and technical responsibility for the study, while SRI, as testing contractor, is responsible for selecting and administering measures both of day care classroom processes and children's development.

The NDCS is being conducted in three phases. Phase I, from July 1974 to September 1975, was a period of planning and implementation. During this period AAI, with the advice and approval of OCD, designed Phase II and formulated criteria for selection of sites and centers. Data collected through telephone interviews with licensed centers in 17 cities resulted in the selection of Atlanta, Detroit and Seattle as the study sites. A total of 64 centers at the three sites (32 in Atlanta; 16 each in Detroit and Seattle) were subsequently selected for participation in Phase II. During Phase I, AAI also developed and tested systems for collecting program and financial information and for assessing parent attitudes. Under a separate contract, SRI selected and field-tested instruments for measuring classroom processes and child outcomes. Phase I results were presented in a three-volume *First Annual Report*[3] prepared by AAI and in a report entitled *Phase II Instruments for the National Day Care Cost-Effects Study*[4] prepared by SRI, in consultation with AAI.

Phase II of the NDCS was a naturalistic study, focusing on centers as they currently operated. Conducted from September 1975 through September 1976, Phase II was designed to build a substantive and methodological base for the study's third phase, which, through experimental manipulation of the policy variables, will provide definitive answers to the policy questions.

Phase II of the NDCS had three objectives:

- To describe fully the natural characteristics of the 64 study centers;
- To test and revise the instrument battery and analytic techniques to be used in Phase III; and
- To further refine the design and implementation plan for Phase III, based on these descriptive and methodological results.

During Phase II, detailed descriptions of all aspects of the 64 study centers were compiled. AAI collected information on program operations, costs, staffing, enrollments, caregiver backgrounds, parent satisfaction and parent use of center services. SRI observed children and caregivers, and administered a test battery to children. All data were integrated into central data files at AAI.

AAI's analytic efforts in Phase II focused on developing and refining a set of theoretically sound, data-based hypotheses about the complex relations among regulatable center characteristics, center processes, costs, and outcomes for children and parents. Phase II descriptive and methodological findings were used to design Phase III, to select a final set of instruments, and to choose sites and centers. All of these Phase II tasks are described in this report, and are detailed extensively in a separate NDCS *Phase II*

Research Report[5] to be published in October 1977.

The descriptive effort of Phase II was necessary because federal day care policy operates in an extremely complex environment. Regulations regarding staff/child ratio, group size, staff qualifications and other center characteristics interact with local regulatory policies and funding mechanisms, with center goals, with demographic characteristics of children and caregivers, with parental expectations, and with a host of other factors in shaping the impact of day care on children and their families. In the absence of prior large-scale research on effects of variations in day care arrangements, it was not feasible at the outset of the study to design an experiment that took full account of the complexities of the day care world. Therefore, instead of proceeding directly from design and instrument development (Phase I) to the experimental study (Phase III), OCD mandated an intervening descriptive phase in which center characteristics would not be altered but rather would be studied as they occur naturally (Phase II).

Phase III, extending from October 1976 through January 1978, is a controlled study designed to test hypotheses based on the results of Phase II. It has two components.

The first component is an experiment being conducted in 49 of the Phase II centers, which were divided into three groups. Group I had relatively low staff/child ratios* in Phase II. These ratios were increased during Phase III by the addition of new staff. Group II had staff/child ratios similar to Group I and was left unchanged. Group III had relatively high staff/child ratios and was also left unchanged. The effects of the ratio increase in Group I centers are being assessed by comparing their classroom processes and child outcomes with those of Group II and III centers.

The second component of the Phase III design is an eight-center, 28-classroom experiment in which children have been randomly assigned, within centers, to classrooms that differ systematically in staff qualifications and staff/child ratios. Again, classroom processes and child outcomes will be compared to assess effects of qualifications and ratios.

Data collection for Phase III will be completed by July 1977. From July through December 1977, Phase III data will be analyzed. The major study questions will be addressed in a policy-focused public report to appear in January 1978. The final technical report will be available for distribution later in 1978.

Parallel with the three-phase investigation outlined above, the NDCS is also conducting a national study of the supply of day care. A telephone survey is being used to compile data on enrollments, staffing, costs and other characteristics of centers. The Supply Study is based on a national probability sample of 3,000 centers, stratified by state. Data will be used to provide a profile of center day care available both across the nation and within states. In addition, the data will aid in interpreting and generalizing the results of the Phase II naturalistic study and the Phase III experimental study. It will allow the NDCS to estimate the national implications of cost-effects findings drawn from Phase II and Phase III. This supporting study began in July 1975 and will be completed in August 1977. Results will be presented in a report to be published in November 1977.

When the research design for the NDCS was first developed in 1974, the FIDCR did not include staff/child ratio requirements for children under three years of age. The Title XX addition of a staffing requirement for infants and toddlers prompted the expansion of the NDCS to include a major descriptive

---

**In this volume, as elsewhere in NDCS documents, "low" staff/child ratio means more children per staff member, and "high" ratio fewer children per staff member; that is, 1:8 is lower than 1:5 and higher than 1:10.*

study of day care arrangements for this age group. Data are being gathered as part of the Supply Study survey and through a series of on-site interviews and observations of centers caring for children younger than three. This study will also be completed in August 1977, and the results will be published concurrently with the final NDCS reports.

ORGANIZATION OF THE SECOND ANNUAL REPORT

The first chapter of this report outlines the design of Phase II of the NDCS, providing a framework for the detailed chapters that follow. Chapters 2 through 4 examine in sequence the major independent and dependent variables of the study. The final chapter presents the Phase III design. Specific chapter contents are as follows:

Chapter 1: Phase II Design and Sample. Centers were not experimentally altered during Phase II but were carefully selected to provide maximum information about the major policy variables — staff/child ratio, group size and staff qualifications. This chapter describes the design of this naturalistic study and the sample of centers and children, focusing on the extent to which the design was successfully implemented.

Chapter 2: FIDCR Policy Variables. Chapter 2 explores alternative methods of measuring the three major policy variables to determine the best measurement approach for research and policy purposes. The chapter covers measurement of staff qualifications, particularly formal education, training in the area of child care and previous day care-related work experience. It also examines different ways of measuring staff/child ratio and group size, and presents Phase II descriptive data useful to the regulator in choosing among these methods. Finally, it discusses other currently or potentially regulatable center characteristics, including (a) center space, equipment and materials, (b) center philosophy and curriculum, (c) availability of specialists and supplementary services to children and families, (d) opportunities for parent involvement in center operations, and (e) continuity of care — the stability of the child's human environment.

Chapter 3: Processes and Outcomes. Chapter 3 outlines the NDCS approach to the measurement of processes in the day care classroom and of outcomes for children and families. The chapter describes two systems used for direct naturalistic observation of children and caregivers, a battery of standardized tests and rating scales used to measure cognitive skills and aspects of social development, and an interview used to assess parent involvement and parent satisfaction with day care. The psychometric properties of the tests and rating scales are reviewed, and analytic and interpretive issues related to naturalistic observations are discussed. Preliminary descriptive findings are reported for each instrument.

Chapter 4: NDCS Financial Analysis. Chapter 4 describes the procedures used to collect financial information from the 64 Phase II centers and presents a preliminary descriptive analysis of day care cost and income patterns. The chapter also describes the quantitative financial model to be subjected to statistical testing and used during Phase III to forecast the impact of variations in regulatable center characteristics on cost per child and other indicators of centers' financial status.

Chapter 5: Phase III Design. Chapter 5 outlines the two-part experiment now underway to measure the costs and effects associated with different configurations of the policy variables, particularly staff/child ratio and caregiver education. Planning and implementation of the two Phase III experimental components are described; the common policy goals and complementary research strengths of the two components are highlighted.

Appendices: The report has four appendices. Appendix A deals with a series of technical issues in the analysis of the effects of day care. Appendix B briefly describes the Supply Study. Appendix C describes the infant care study. Appendix D lists staff and consultants of the NDCS and briefly describes the role of those consultants, including a newly formed task force of black consultants.

PURPOSES AND LIMITATIONS OF THE SECOND ANNUAL REPORT

Most findings of the National Day Care Study relevant to government policy will be based on Phase III data. The purposes of Phase II were to refine data collection systems and instruments, to describe day care as it occurred naturally in the 64 study centers, and to develop a powerful, workable Phase III design based on this detailed knowledge of the study sites and centers. This report, consequently, presents primarily descriptive and methodological findings and does not, in general, deal with the extension of these to conclusions regarding policy.

In Phase II, associative relationships among policy variables, center processes, and child and parent outcomes were explored in order to develop hypotheses for Phase III. All of these hypotheses require careful testing before any can be accepted or rejected as statements about the effects of federal policy on day care. It would be irresponsible to report these hypotheses here, since they could easily be mistaken for policy conclusions rather than refined statements of policy questions.

Nevertheless, there are some policy implications of the Phase II measurement procedures; for example, NDCS findings about the properties of various methods of measuring staff/child ratio are relevant for monitoring compliance with ratio requirements. Policy implications of this kind are discussed in this report.

Chapter 1: PHASE II DESIGN AND SAMPLE

Phase II of the NDCS provided the substantive and methodological foundation needed to build the final Phase III research and analytic plan. During Phase II, relationships among center characteristics were explored as they existed naturally. This analysis identified the center characteristics to be experimentally manipulated during Phase III in order to systematically illuminate their effects on child and family outcomes and on costs. This chapter discusses the rationale for the Phase II research design, describes the sample of 64 centers selected for the natural study and presents the implications of Phase II results for the experimental design of Phase III.

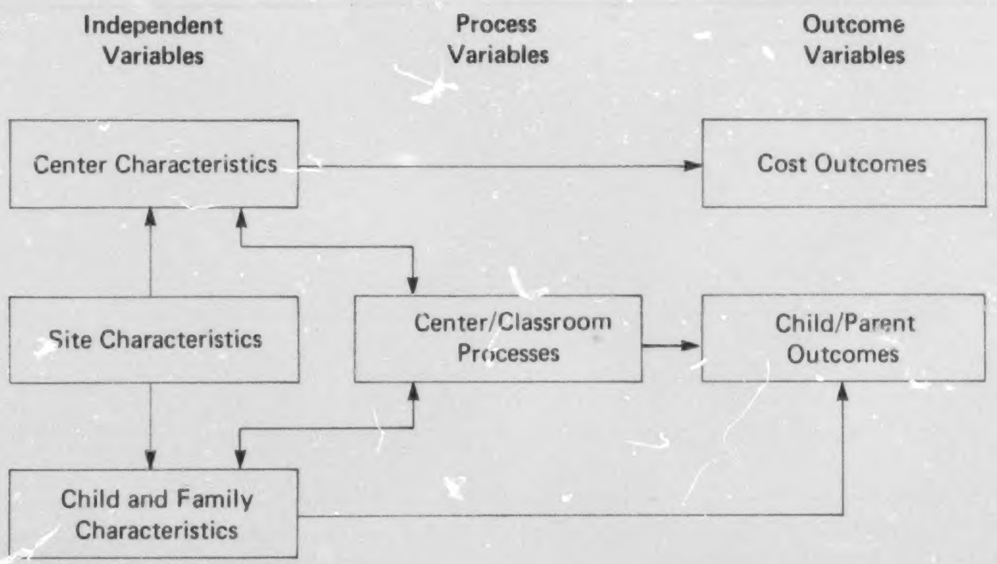
PHASE II DESIGN

Underlying the basic design of both Phases II and III is a conceptual model of center day care illustrated in simplified form in Figure 1-1. The figure reflects the following general assumptions:

- characteristics of centers, children and families vary from site to site;
- child and family characteristics and center characteristics influence center and classroom processes, particularly interactions between caregivers* and children;
- center and classroom processes affect children's social, emotional and cognitive development, and parents' satisfaction with their day care arrangements;
- child and family characteristics also affect outcome measures directly, without mediation of classroom processes; family background, for example, has an effect on a child's social development that is independent of any day care influence;
- center characteristics – enrollment, staff/child ratio, staff qualifications and many other factors – determine the per-child cost of day care.

Figure 1-1

NDCS Model of Center Day Care


Independent Variables
Process Variables
Outcome Variables
Center Characteristics
Cost Outcomes
Site Characteristics
Center/Classroom Processes
Child/Parent Outcomes
Child and Family Characteristics

In the NDCS, "caregivers" refers to classroom staff – teachers and aides.

The objective of the NDCS is to estimate the causal relationships in this model and to use the results to evaluate the costs and effects of alternative configurations of day care center characteristics.

Of primary interest to the NDCS are those center characteristics that are, or could be, manipulated by federal day care regulations. As noted earlier, three of these characteristics have been termed "policy variables" – staff/child ratio, group size and caregiver qualifications. Although these variables have received primary attention during Phase II, other regulatable characteristics of centers have also been studied – space, equipment and materials, center philosophy and curriculum, availability of specialists and supplementary services, opportunities for parental involvement, and continuity of care. (These other variables are discussed in Chapter 2.)

The Phase II research design focused on selection of a sample of centers which could be used to estimate effects of regulatable center characteristics. No center characteristics were experimentally manipulated during Phase II. The research design objective was to provide a set of rules for center selection which, in the absence of experimental manipulation, would maximize the usefulness of the sample in disentangling the effects of regulatable variables.

The most effective approach for assessing the impact of a particular policy variable, without deliberate manipulation, was to choose centers which were alike in all but that one variable. For example, if one group of centers had high staff/child ratios and another group had low ratios but the two groups were otherwise alike, any differences in child outcomes between the two groups could be attributed to the difference in ratios. A similar approach could be used to control the variability in caregiver qualifications and group size. Some centers were selected for caregivers with above-average formal education; other centers were selected for caregivers with below-average formal education.* Some centers were chosen for their large group sizes (above-average number of children per classroom), others for their small group sizes.

If staff/child ratio, caregiver education and group size are classified for each center as high or low, day care centers can be divided into eight possible categories (cells), as shown in Figure 1-2. During the sampling process for Phase II, all potential study centers were sorted into the eight classification cells, and final selections were then made from within each cell.

Technically, the Phase II research design is a three-variable, two-level (high/low) *factorial design*. This design maximizes the variation across centers in each of the policy variables by including both high and low levels for each variable. The design also ensures the independence of the policy variables by including centers which are alike in all but one variable at a time. The high variability and independence of the policy variables across centers in the research sample gives the design considerable statistical power to disentangle the effects of the policy variables, operating singly or in combination.

The design was replicated at each of the three study sites. Four centers were selected per cell in Atlanta (32 centers) and two centers per cell in Detroit and Seattle (16 centers each). The three sites were selected to maximize diversity of geographic, demographic and economic environments in which centers operate, as well as diversity of local funding mechanisms and regulations. This diversity was deliberately sought to determine whether or not the impact of federal regulation is affected by local conditions and, if

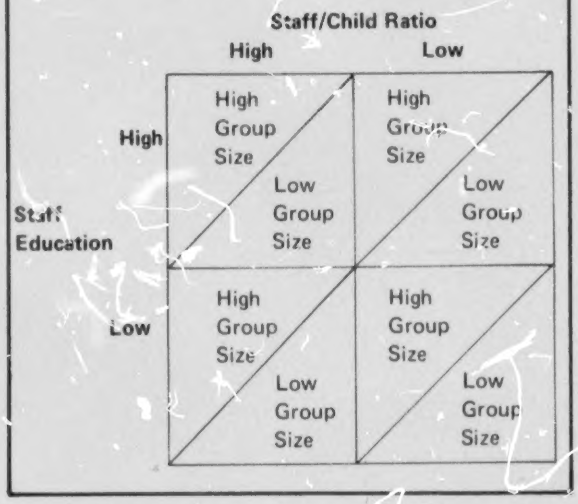
---

**Formal education is only one component of caregivers' qualifications. Day care-related experience and training, including on-the-job training, are also important components. Education was given primary emphasis in center selection, but diversity of training and experience was also sought in selecting the Phase II research sample.*

13

Figure 1-2

The Phase II Eight-Cell Design



so, in what ways. In addition, diversity of sites aids in assessing generality of the NDCS findings. If effects of the policy variables are similar across very different sites, there will be good reason to believe that research findings hold generally and provide a sound basis for national policy. If effects differ across sites, examining characteristics of sites will provide a way of better defining the proper scope and limits of federal regulation.

Diversity was also sought among the individual centers and the children and families they served. Within the framework of the eight-cell design, centers were selected to vary as much as possible in regulatable and non-regulatable characteristics other than the policy variables themselves. For example, centers serving substantial numbers of both black and white children were recruited — including integrated centers and those serving predominantly black or white groups of children. Similarly, centers were sought that served both low- and middle-income families and therefore included substantial numbers of children supported by public subsidy and children supported by parent fees. Finally, an effort was made to recruit centers that operated under a variety of auspices and drew their funds from different sources.

Although the Phase II design systematically controlled the variation of the policy variables and their independence from each other, it was not possible to deal with the other potential causal factors in so precise a fashion. High variability and independence were sought for many of the other variables which potentially could influence child and family outcomes and day care costs — center size, profit/non-profit status, FFP/non-FFP status,* racial and socioeconomic composition of enrollment, and the day care-related training and experience of caregivers. The list of explanatory variables — potential determinants of the cost and quality of care — is long; it was not feasible within the sample of study centers to obtain an appropriate range of values for every possible variable. One goal of Phase II was to determine the degree to which effects of the policy variables were confounded (entangled) with effects of these additional explanatory variables. Once identified, this confounding could then be eliminated by experimental manipulation of center characteristics at the start of Phase III.

Diversity of the Phase II sample has been stressed up to this point. It should also be noted that sample selection was constrained in a number of ways, as indicated by the following selection criteria used to determine eligibility of centers for inclusion in the research sample[1]:

- *Providing year-round full-day care:* To be eligible for participation in the study, it was required that centers be open at least seven hours per day, five days per week and 10 months per year. Only full-time year-round centers offer day care arrange-

---

**The term FFP (federal financial participation) is used to denote centers which are reimbursed for day care services under Title XX of the Social Security Act. Such centers are expected to comply with the FIDCR, in addition to state day care licensing standards.*

ments which satisfy one intent of federal day care appropriations under Title XX – promoting economic self-sufficiency by freeing parents for training and work.

- *In operation at least one year:* To increase the probability that centers would continue in operation throughout Phases II and III, and to avoid studying non-recurring start-up behavior, centers were required to have been in operation for at least one year at the time they were selected.
- *Located in urban areas:* Realistic limitations on the costs of testing children and collecting program and financial information from centers required that the NDCS sample be located in geographical clusters. Only centers in or near urban areas satisfied this criterion. Since a large percentage of low-income families are also located in such areas, this selection rule excluded only a small proportion of the policy-relevant population.
- *Serving three- and four-year-old, English-speaking children:* The research sample for Phase II and III was limited to children who were three and four years of age at the start of each testing period. Since three- and four-year-olds constitute more than half of the day care population, they represented a high priority study group. Children from non-English-speaking families were not included in the research sample. Adequate test batteries for non-English-speaking children did not exist; in addition, the non-English-speaking group constitutes a small percentage of the day care population. (The infant and toddler substudy was added in 1976; infants and toddlers are not included in the cost-effects study.)
- *Having an adequate sample of full-time children:* The turnover rate for children in day care centers was expected to be high. To ensure that start-of-year and end-of-year test and observation data would be available for an adequate number of children, centers were included in the sample only if they had fifteen or more three- and four-year-old children enrolled on a full-time basis.

Some of the selection criteria used in the NDCS were dictated by research considerations – cost of data collection, adequacy and stability of the sample, and feasibility of measurement. Others were imposed to increase representation of the policy-relevant population – low-income children receiving publicly subsidized, work-related care. Data are available from the Supply Study on the percentage of centers and children nationally which meet the selection criteria used in choosing the NDCS sample.[2] About 66 percent of the day care centers, 69 percent of the total enrollment and 71 percent of the publicly subsidized enrollment in the United States are found in or near major urban areas. Only 44 percent of U.S. centers meet the criteria on minimum center size, but these account for 64 percent of total center enrollment and 71 percent of both publicly subsidized enrollment and full-time three- and four-year-old enrollment. Thus, a significant proportion of the U.S. day care center population, and an even larger proportion of the policy-relevant population, is represented in the NDCS research sample.

DESCRIPTION OF THE PHASE II SAMPLE

The research sample for Phase II consisted of 64 day care centers – 32 in Atlanta and 16 each in Detroit and Seattle. At the start of Phase II there were approximately 4500 children enrolled in the study centers. About 1800 of these were three- and four-year-olds enrolled on a full-time basis; the remaining children were infants, toddlers or school-age children or were three- and four-year-olds enrolled only part-time. About 1500 staff were employed in the study centers, half of them in the classroom as teachers or aides.

Three- and four-year-old children enrolled in study centers for at least seven hours per day, four days per week were designated as "target children" and constituted the potential sample of children for Phase II analysis. Classes in which there were at least ten such children or in which they represented a majority were designated as "target classes."[3] At the start of Phase II (fall 1975), only target children in target classes were eligible to be tested or observed. Parental permissions were obtained to test and observe 1384 children, and start-of-year outcome measures were obtained for 1328 children. To increase the number of children for whom some Phase II, as well as Phase III, outcome measures would be available, the group to be tested and observed was expanded in spring 1976 to include three-year-old children in non-target classes. End-of-year outcome measures were obtained for 1502 children. Table 1-1 provides detailed information on the Phase II sample of children.

Characteristics of Centers. At the start of Phase II, enrollment in study centers ranged from 16 to 150 children. Average enrollment rose from 70 children per center in October to 80 children in April — reflecting the fact that full enrollment had not been reached when fall data were collected. Forty-nine study centers were operated on a non-profit basis; the remaining 15 were operated as either proprietorships or for-profit corporations. Of the non-profit centers, 12 were run by public agencies (nine by the Atlanta public school system and three by Head Start), 19 were sponsored by religious organizations and 13 were supported by voluntary agencies; the remaining centers were independent. For almost half of the study centers, government funding was the largest single source of income; for the remaining centers parent fees were the primary source of revenue.[4] Most of the centers in Atlanta (22 of 32) relied on government funds as their primary source of income. Most of the centers in Detroit (13 of 16) and Seattle (10 of 16) relied primarily on parent fees.

Characteristics of Staff. Virtually all classroom staff and about three-fourths of the non-classroom staff in the Phase II sample were female. About 60 percent of the sample staff were black, 38 percent white and less than two percent were of other racial groups. The percentage working full-time was higher for classroom staff than for non-classroom staff, but in both cases less than half worked 30 hours or more per week. Classroom staff tended to be young (60% were under 30), about half had children of their own, and one-third were principal breadwinners for their families. Center averages in years of education for classroom staff ranged from 10.3 to 16.9 years; the typical caregiver had completed high school and had received some post-secondary education. Center averages in years of day care-related experience of caregivers ranged from less than one year to 9.6 years; the typical caregiver had between two and four years of experience. Caregivers in the Atlanta sample had slightly more previous experience than caregivers in Detroit and Seattle, but educational attainment was comparable across the three sites.

Characteristics of Target Children and Families. About 64 percent of the target children were black, a third were white and the remaining three percent of other racial origin. Half the target families were single-parent households, although many of these households contained other adults — grandparents, uncles, aunts or older siblings. There were 3.8 persons in the average child's household. Three-fourths of all mothers were employed either full- or part-time and the remainder were in school or a job-training program. About 90 percent of the fathers present in the home were employed. Average education was 13 years for mothers and 14 years for fathers. Employment was the primary source of income for three families out of four.* More than a quarter of the sample families received some welfare assistance, but welfare was the primary source of income for

---

**Many of these families received some type of public child care subsidy, usually in the form of grants to providers under Title XX. This provides funds to purchase child care for mothers who would otherwise be or who were formerly welfare recipients. On the average, study centers received approximately 40 percent of their funds from government sources.*

Table 1-1
NDCS Phase III Sample of Children

	Number of Children	
	Fall 1975	Spring 1976
Full-Time Three- and Four-Year-Olds	1819	1863
Eligible to be Tested and Observed[a]	1608	1753
Parent Permission Obtained	1384	1644
Parent Permission Refused	27	35
Unable to Contact Parents	197[b]	74
Tested or Observed or Both	1328	1502[c]
Tested Only	1260	1437
Observed Only	1244	1350
Tested and Observed	1176	1285

	Number of Children
Tested in Fall and Spring	987
Observed in Fall and Spring	920
Tested and Observed, Fall and Spring	883

[a] Only target children in target classes were eligible in the fall; to increase the sample of longitudinal data for Phase III, the eligible group was expanded in spring 1976 to include full-time three-year-old children in non-target classes.

[b] A much higher number of parents were not contacted in fall 1975 than in spring 1976. In Atlanta, where 181 of the 197 failures to make contact occurred, many parents did not have phones and their children were picked up by center-provided buses. A major storm delayed the process of contacting these parents at their homes for a crucial week just prior to the fall testing period. The percentage of permissions secured from parents of eligible children rose from 86% in fall 1975 to 94% in spring 1976.

[c] In spring 1976, observation and/or test data were collected for 91% of the children for whom parental permission had been obtained – down from 96% in fall 1975. Between the time parental permissions were obtained and the time testing and observations began, the imposition of more stringent day care eligibility requirements by the state of Georgia (under Title XX) forced a significant number of previously subsidized children out of study centers. This sudden, one-time attrition of children in Atlanta was responsible for the lower testing/observation rate in spring 1976.

less than one family out of six. The remaining nine percent cited alimony, disability benefits, social security, pensions and other retirement benefits as their primary source of income. Family income was evenly divided across three categories — under $6000, $6000 to $15,000, and over $15,000. Measures of SES were highest for Detroit families and lowest for Atlanta families. Within all sites, however, the range of socioeconomic status was quite broad.

*Staff/Child Ratio and Group Size.** The average staff/child ratio for the sample of target classrooms was 1:7.3. Across the 64 sample centers, staff/child ratios ranged from 1:2.5 to 1:23.3, although the majority of centers had ratios between 1:5 and 1:11. Three- and four-year-olds were generally grouped separately within centers, but staffing ratios did not differ appreciably for the two age groups. Group size is a measure of the number of children under the supervision of one caregiver or a team of caregivers within a single classroom. Centers were typically organized with a single group of children per room. In centers organized on an open-classroom basis, however, several groups of children were often found in a single large room. The average group size in the sample was 18.6 children. While center averages ranged from 6.9 to 56.5, most centers had group sizes between 10 and 25.

Average ratios and group sizes were comparable across sites. However, the five centers with ratios higher than 1:5 (fewer than five children per caregiver) were in Atlanta, while the only two centers with group sizes smaller than ten were in Seattle.

EVALUATION OF THE SAMPLE

As noted earlier, the ability of the Phase II study to provide information about effects of the major policy variables — staff/child ratio, group size and staff qualifications — depended on three conditions: (1) adequate variation within each policy variable; (2) absence of confounding relationships among the policy variables; and (3) absence of confounding relationships between the policy variables and other potential determinants of processes and outcomes, such as racial and socioeconomic characteristics of children being observed and tested. This section demonstrates that the first two of these conditions were met. Despite the careful procedures used in selecting study centers, however, variations in the socioeconomic characteristics of children across centers were linked in the Phase II sample to variations in both staff/child ratio and staff experience. This confounding had important implications for both analysis of Phase II data and design of Phase III.

Variation in Staff/Child Ratio, Group Size and Staff Qualifications

The section of this chapter describing the Phase II centers showed that the sample included broad ranges of staff/child ratio, group size, staff education and staff experience, although many of the centers were concentrated within narrower ranges of ratio, group size and education. In judging how well these distributions met the needs of the Phase II study, three questions must be considered: (1) Did the policy variables show enough variation to give the study adequate statistical power to detect their effects? (2) Were all levels of the policy variables that are relevant for federal regulation represented by at least some centers in the sample? (3) Were most or all levels of the policy variables that exist among centers nationally represented by at least some centers in the sample? Because statistical power – the ability of the study to detect effects – depends on independence of the policy variables as well as their variation, the first question is treated after relationships among the policy variables have been discussed. The second and third questions are discussed immediately below.

---

* *Data presented here on staff/child ratio and group size are time-weighted averages of observations of all adults and children (including volunteer staff and part-time caregivers and children) in target classrooms during morning hours in May 1976. As pointed out in Chapter 2, staff/child ratios are significantly affected by the method used to gather data (observation vs. examination of center records), by decisions about which staff and children to include, by the time of day during which ratios are assessed, and by the method used to compute averages. For reasons outlined in Chapter 2, the NDCS has chosen time-weighted averages of observed morning ratios as the measure most likely to show an impact on children's development. However, it should be recognized that this method of measuring and computing ratios tends to produce relatively high values of ratio (relatively few children per caregiver). Thus the ratio figures shown here are higher than those calculated by other, more widely used methods.*

The Phase II centers were chosen to meet specific design requirements. They were not sampled randomly from the national population of day care centers; therefore, they could not be expected to exhibit proportional representation of all the different characteristics of centers nationwide. In fact, the study's selection criteria guaranteed that the sample would include more than the national proportion of certain kinds of centers (e.g., large centers, centers serving three- and four-year-olds, publicly funded centers) and less of others. Appendix B contains a comparison of the Phase II sample and the nationally representative sample drawn for the Supply Study. The comparison highlights some of these differences and shows how they arose from the differing goals and selection criteria of the two studies.

In order to provide a data base for formulation of hypotheses about effects of different levels of the policy variables, it was not necessary that the Phase II sample show proportional representation of the different levels that prevail nationally. It was, however, necessary that the sample include centers at every level that might be mandated by foreseeable federal regulations. In addition, it was desirable for comparative purposes that the sample include some centers having characteristics lying outside the range of likely federal regulations, but present among centers nationally.

The Phase II sample included centers with average staff/child ratios* near or above the FIDCR requirements of 1:5 for three-year-olds and 1:7 for four-year-olds. The sample also included many centers operating substantially below FIDCR levels. Similarly, the sample included centers near or below the FIDCR group size maxima of 15 for three-year-olds and 20 for four-year-olds, as well as many centers above FIDCR levels. Thus the sample allowed the NDCS to distinguish effects of those levels of ratio and group size currently mandated by the federal government from the effects of a range of less stringent levels.

The range of ratios in the sample also permitted the NDCS to distinguish any significant effects associated with many current state regulations. In 1976, four states imposed ratios from 1:5 to 1:7 for three- to four-year-olds. Thirty-one states mandated ratios from 1:8 to 1:11, and fourteen allowed ratios lower than 1:11 (the lowest being 1:20). Two states had no ratio requirements.[5] The Phase II sample, with its concentration in the range from 1:5 to 1:11, allowed detailed discrimination of effects within the range of most state regulations, while also permitting a more global comparison of effects associated with very high and low ratios.

While the FIDCR do not set specific requirements for staff education or experience, the ranges of these variables across the Phase II centers (10.3 to 16.9 years of education; eight months to 9.6 years of experience) were so broad as to include any minimum level of staff education or experience that might foreseeably be mandated by the federal government. More relevant for current federal policy is the fact that the sample included a substantial proportion of caregivers with less than a high school education and

---

**Again, it must be remembered that these ratios were derived by methods that tend to yield high values – relatively few children per caregiver. Alternative methods of measuring and computing ratios are discussed in Chapter 2.*

with little or no previous child care experience. The FIDCR requires centers to provide employment opportunities for low-income people, who often lack formal education and work experience. Thus, determining the effectiveness of caregivers who lack formal qualifications is directly relevant for evaluating this aspect of current policy.

Relationships Among the Policy Variables

In the Phase II sample, centers that had high staff/child ratios (relatively few children per staff member) also tended to maintain smaller groups than low-ratio centers. The relationship between the two variables, however, was not strong. Aside from this modest association between ratio and group size, the policy variables were essentially unrelated to each other. Table 1-2 shows the relationships among staff/child ratio, group size, staff education and staff experience. All relationships are presented in the form of correlation coefficients. Thus the Phase I selection process largely achieved its goal of recruiting a Phase II sample of centers in which the policy variables were independent of one another. As pointed out earlier, the absence of strong relationships among the policy variables is extremely important because it maximizes the ability of the study to identify effects of the individual policy variables on center processes and child outcomes.

Table 1-2
Correlations Among the Policy Variables

	Group Size	Staff Education	Staff Experience
Staff/Child Ratio	−.24	+.02	+.03
Group Size		+.06	−.08
Staff Education			+.06

NOTE: Each decimal fraction is a correlation coefficient, or measure of association between the variables. Correlation coefficients can vary from a maximum of +1.0, indicating the strongest direct relationship, through 0.0, indicating no relationship at all, to −1.0, indicating the strongest inverse relationship.

The variables shown in the table are defined as follows:

Staff/Child Ratio = average ratio observed in all target classes in each center during the morning hours in May 1976.

Group Size = average number of children observed in all target classes in each center during the morning hours in May 1976.

Staff Education = average number of years of education for directors, target classroom teachers and aides in each center.

Staff Experience = average number of years of child care experience for directors, target classroom teachers and aides in each center.

Relationships Between the Policy Variables and Other Center Characteristics

The most important challenges to Phase II analysis and the most important lessons learned for Phase III design arose from the relationships between the policy variables and other center characteristics. Chief among these characteristics were differences among centers in the kinds of children and families served.

Many previous studies have shown that children's socioeconomic status (SES) and race are often related to their test scores and behavior. While SES and race are not the only important measures of child background characteristics, they tend to be associated with other, less easily measurable variables, such as home atmosphere and parental childrearing practices, which are also potential determinants of test scores and behaviors.* Thus, if SES and race were strongly related to any of the policy variables — that is, if children of one SES or racial group were predominantly found in centers with high or low levels of one of the policy variables — effects of these background characteristics could mask policy variable effects or create apparent effects where none existed.

As Table 1-3 shows, the staff/child ratios maintained by centers were related to the SES of children served, as indexed by center-average levels of family income and educational attainment of mothers. Children from low-income families and children whose mothers had relatively low levels of education tended to be found in centers with high staff/child ratios (few children per caregiver). As would be expected, the two SES measures — family income and mothers' education — were closely related to each other. The relationship between SES and ratio was linked to federal funding: low-income children were found predominantly in publicly funded centers, which maintain high ratios, presumably because of FIDCR ratio requirements.

Staff experience was also related to family income and public funding in the sample. Low-income children were exposed to more experienced staff, and more experienced staff were found in centers with high proportions of public funding. There was also a weak tendency for children of less-educated mothers to be found in centers with more experienced staff. These relationships are consistent with the FIDCR mandate that, in federally funded centers, "Educational activities must be under the supervision and direction of a staff member trained *or experienced* in child growth and development."[6] (Emphasis added.)

Race was almost entirely unrelated to the policy variables. Although black children tended to be found in centers with high ratios and less-educated but more experienced staff, these tendencies were extremely weak. Thus effects of race on children's test scores and behavior did not cloud the analysis of policy variable effects. One reason for the low degree of confounding between race and the policy variables was that a deliberate effort was made in Phase I to recruit centers serving middle-class black families. Because of this effort, race and SES were only moderately related in the NDCS sample, as were race and public funding of centers. Thus black children were found both in publicly funded centers with high staff/child ratios and experienced staff and in parent-fee centers with lower ratios and less experienced staff.

Implications for Phase II Analysis and Phase III Design

In order to draw definitive conclusions about effects of each of the policy variables, it is

---

In recognition of the fact that large variations in home environment exist within racial and SES groups, the NDCS incorporated interview measures of home atmosphere and parental practices in its Phase III data collection plans. The study also used a direct measure of one important initial characteristic of children (the child's entering score on the Preschool Inventory, described in Chapter 3), rather than race and/or SES, as an independent variable in the Phase III design.

Table 1-3
Correlations Between the Policy Variables and Other Center Characteristics

	Family Income	Mothers' Education	Percent Black Children	Percent Federal Funding
Staff/Child Ratio	−.34	−.21	+.11	+.47
Group Size	+.01	−.07	−.07	+.03
Staff Education	+.09	+.14	−.15	−.05
Staff Experience	−.36	−.18	+.19	+.31

NOTE: Decimal fractions shown in the table are correlation coefficients, exlained in the note to Table 1-2. Definitions of the policy variables are in the same note to Table 1-2. Definitions of other center characteristics are as follows:

Family Income = average annual income of families of children served in each center.

Mothers' Education = average number of years of education for mothers of children served in each center.

Percent Black Children = proportion of all children enrolled in each center who were black.

Percent Federal Funding = proportion of the total center budget that is funded by the federal government.

necessary to separate each variable from all other potential determinants of processes and outcomes. This is best accomplished by an experimental design in which variables of interest are systematically manipulated, while all others are held constant. A manipulation of this sort was implemented in Phase III. When such manipulation is not possible, as in a naturalistic study like Phase II, the separation of variables can be partially achieved by careful sample selection. Even with careful selection, however, complete separation of explanatory variables is rarely possible. As a result, statistical adjustments must be used to compensate for confounding relationships among important independent variables.

The Phase I center selection process resulted in a group of Phase II centers in which staff/child ratio, group size, staff education and staff experience varied widely, with most centers concentrated in the ranges most relevant to policy. In addition, the policy variables were almost completely independent of one another, with the exception of a modest relationship between group size and ratio. The deliberate selection of a sample characterized by wide variability and independence of the policy variables gives the study considerable ability to detect effects of the policy variables on processes and outcomes. This was demonstrated in a series of technical investigations, reported in Appendix A and discussed in detail in the *Phase II Research Report*. Essentially these investigations showed that, if the policy variables had large effects, those effects would be virtually certain to be detected. If they had effects of moderate size, they would be very likely to be detected. Effects small enough to escape detection would be unlikely to have much practical significance.[7]

Moderate relationships existed in the Phase II sample between child SES and both staff/child ratio and staff experience. Much weaker relationships existed between race, on the one hand, and ratio, staff education and staff experience on the other. In Phase II, a standard statistical approach of covariance analysis was used to disentangle the effects of ratio and experience from those of child SES and race. This approach is discussed further in Appendix A.

Phase III has two major components, one of which was designed to eliminate confounding of staff/child ratio and child background characteristics. This experiment, described in detail in Chapter 5, will involve manipulating ratios in one set of centers, while leaving ratios unchanged in a second closely matched group of centers.* Staff/child ratio was singled out for manipulation since the ratio requirement has been the most controversial of all FIDCR provisions because of its potential cost implications. Staff experience was balanced across control and treatment groups in the Phase III design, so that its effects could again be assessed without manipulation as in Phase II. Staff education and group size remained unconfounded with other policy variables and child background characteristics; thus their effects also can be reassessed in Phase III.

In sum, the Phase II sample provided an adequate basis for formulating hypotheses about effects of staff/child ratio, group size, and staff education and experience, because these FIDCR-regulated center characteristics showed sufficient variability and independence among the Phase II centers to permit disentanglement of their effects. However, since ratio and experience were related to the socioeconomic characteristics of children, statistical adjustments were required to separate effects of these two policy variables from those of the children's backgrounds. One component of the Phase III experiment has been designed to reduce the need for such adjustments in order to provide the clearest possible information on effects of ratio. The effects of the other policy variables will be assessed by methods like those used in Phase II.

**For reasons explained in Chapter 5, a direct measure of one aspect of children's development – their entering scores on a standardized test – was used for purposes of matching the manipulated and control centers. This approach had the effect of near-matching race and SES as well.*

Chapter 2: FIDCR POLICY VARIABLES

INTRODUCTION

During Phase II, the NDCS explored alternative ways of measuring the major policy variables – staff/child ratio, group size and staff qualifications – as well as other potentially regulatable center characteristics. These explorations accomplished four objectives:

- clarification of measurement issues relevant to both the researcher and the policymaker;
- selection and refinement of measurement techniques to be used in Phase III;
- description of the policy-relevant characteristics of Phase II centers; and
- identification of implications of those characteristics for regulation and monitoring.

This chapter addresses issues pertaining to measurement of regulatable center characteristics – specifically, staff qualifications, staff/child ratio and group size. The primary objective of the chapter is to present the measurement techniques and definitions to be used in Phase III research. Many Phase II findings should be useful, however, in clarifying the language of day care regulations and in refining techniques for monitoring regulatory compliance. These policy and administrative implications of Phase II results are also addressed here. The chapter begins with a discussion of goals and rationales underlying various provisions of the FIDCR and an analysis of criteria that operational definitions of day care characteristics must meet for regulatory and research purposes.

The Goals and Rationale of the FIDCR

The provisions of the 1968 FIDCR, as amended by Title XX and outlined in more detail in the NDCS *First Annual Report,*[1] cover a variety of center characteristics, including:

- staff/child ratios and group sizes for children of different ages;
- location, licensing, sanitation, safety and general suitability of the center's physical facility;
- educational opportunities, which must be under the supervision of qualified staff. (A separate formal educational program is not required under Title XX.);
- social services to children and families;
- health and nutrition services;
- in-service training of staff;
- parent involvement;
- program administration and coordination; and
- program evaluation.

Although the framers of the FIDCR did not provide detailed, explicit justifications for the various provisions of the regulations, it is apparent that two broad goals were paramount:

- protecting children in publicly funded day care from exposure to potential harm; and
- ensuring that children in publicly funded day care will be exposed to conditions believed to promote all forms of development.

In addition, the FIDCR embody several more specific goals, some of which can be seen as means to the broader ends of preventing harm and promoting development. These goals include:

- guaranteeing a parental voice in center decision-making;
- promoting the general welfare of families through supplemental services;
- ensuring responsibility and accountability on the part of program administrators; and
- promoting employment of low-income people in day care centers.

The two broad goals of preventing harm and promoting development lead to regulations that differ in important ways. For example, consider the number of adults required to supervise a given group of children: relatively few might suffice to prevent accidents and keep conflict within bounds, while a larger number might be needed to guide children in educational activities, particularly when such activities require recognition of children's individual needs. *Whether* federal regulations should be aimed at promoting development or merely preventing harm is a philosophical and political question, currently the subject of heated debate. *Which* specific regulations will best accomplish these goals is in part a factual question to which the NDCS will provide answers.

Most state regulations and many provisions of the FIDCR, especially those dealing with safety and sanitation standards, are aimed principally at preventing harm. These FIDCR provisions have been relatively uncontroversial. In contrast, some of the FIDCR provisions aimed at promoting development have been attacked. On one hand, some critics see these provisions as unrealistic or excessively expensive. The most notable example is the staff/child ratio requirement. Debate has been so strong that for Title XX day care, ratios for center-based care have been frozen at September 1975 levels for children under six years, pending further investigation of the appropriateness of the FIDCR ratio provisions.

On the other hand, some critics have argued that the FIDCR are insufficiently oriented toward developmental goals. For example, the provisions for staff qualifications are neither explicit nor stringent enough to control the kinds of background that some believe caregivers should bring to their work.

As indicated previously, the NDCS has focused on the controversial ratio requirement, on the group size requirement and on the less specific requirements regarding staff qualifications. All three of these requirements arise from the same basic assumption about the effects of day care, namely that certain kinds of interactions between staff and children are crucial for preventing harm and promoting development. Both a substantial body of research literature in child development and the weight of practical experience support this assumption. Such interactions cannot be regulated directly, however, since they are influenced by personal characteristics of caregivers and children, by the activities of the moment and by many other factors outside the scope of government control. Consequently, the government has attempted, through its regulatory policies, to establish facilitating conditions for positive interaction. Staff/child ratio and group size as features of classroom social structure are presumed to influence the amount of time, energy and attention that caregivers can devote to children.[2] Training and experience are presumed to foster the caregiver's sensitivity to children's needs and the ability to meet them.

Other center characteristics may also be associated with the events a child encounters daily and with the corresponding impact on both child and family. Personal qualities of the director, staffing patterns and schedules, rates of pay, turnover, physical characteristics of the center (such as space, equipment and materials) and, most obviously, the programmatic elements of the center (developmental emphasis, classroom curriculum and support services for children and parents, for instance) — all these are clearly important to the dynamics of the day care center. Some are currently regulated and others might be candidates for regulation.

Determining whether any of the abovementioned variables shows measurable effects on classroom processes or child outcomes in the face of a great deal of unregulatable variation across centers is a major research goal of the NDCS. But before any assessment of effects could be undertaken, it was necessary to address certain problems of measurement that are built into the FIDCR as currently worded.

Issues of Monitoring and Measurement

In their present form, most provisions of the FIDCR are too imprecise to permit uniform interpretation, monitoring and enforcement.[3] For example, regarding staff qualifications the FIDCR state only that "Education must be under the supervision of a person trained or experienced in child growth and development" and that "The persons providing direct care for children . . . must have had training or demonstrated ability in working with children."[4] Specific educational requirements, levels of experience or types of training are not stated, nor are acceptable mixes of staff with different qualifications. For example, can inexperienced aides provide "direct care" if they are supervised by lead caregivers who "have had training or demonstrated ability in working with children"? Unless supplemented by elaborate monitoring rules, such provisions do not allow a monitor or center to determine whether the center is in compliance.

The ambiguity that confronts the monitor also confronts the researcher. In order to assess the effects of regulated center characteristics, uniform definitions of those characteristics must be used, but the FIDCR do not provide such definitions. The NDCS has responded by examining several different methods of measuring each policy variable.

Measurement Criteria: Alternative measures of the policy variables were tested against three criteria:

- reliability;
- validity; and
- utility in monitoring and regulating.

Reliability and validity are standard research criteria. A measurement technique is reliable if it gives the same results when applied by different people on different occasions. A measurement technique is valid if it actually assesses the quantity or quality it is designed to assess. When applied to measures of regulated characteristics of day care centers, these research criteria have implications for regulation and monitoring. An unreliable technique for measuring staff/child ratio — one that gives different results when applied by different monitors on different occasions -- cannot be used to judge whether a center is in compliance with regulations. Similarly, an invalid technique — one that gives consistent but misleading results — cannot be used. For example, center records would be unreliable measures of ratio if the records were kept inconsistently from day to day. They would be invalid measures if they systematically failed to count certain staff (part-time staff or volunteers) or certain children (part-time children or "drop-ins").

A criterion that has been applied consistently in the NDCS to evaluate alternative definitions and measurement techniques is utility in monitoring and regulation. Measures must be precise and unambiguous, so that compliance can be readily determined by both providers and monitoring agencies. Measures must be flexible enough to suit the wide variety of day care situations which are found in practice — mixtures of part-time and full-time enrollment, mixtures of parent-fee and publicly subsidized children, open classroom arrangements with multiple groups of children in a single room, and hierarchical staffing arrangements of lead teachers, teachers, aides and volunteers. Finally, good definitions and measurement techniques must be feasible to implement. They must not impose excessively expensive and time-consuming record-keeping requirements, and they must not require frequent and/or highly intrusive monitoring of center and classroom operations.

Other Contributions of the NDCS to Regulatory Decisions: There are two additional requirements for regulation that go beyond considerations of monitoring. First, regulations must be feasible for centers to implement, not just for monitoring agencies to enforce. That is, regulations must not be so remote from prevailing practices that the cost of compliance would be prohibitively high or center operations significantly disrupted. Second, regulations must have tangible and important positive effects, in order to justify imposition of penalities for noncompliance. **Measurement of costs and effects associated with different configurations of regulated center characteristics is the task of Phase III of the NDCS. Phase II findings bear only on the validity, reliability and feasibility of implementation of alternate definitions and measurement techniques.**

Findings of Phase II and Phase III should also aid the policymaker in drawing the important distinction between regulations (requirements) and guidelines (recommendations). Regulations, by definition, involve penalties for noncompliance. In the case of the Title XX FIDCR, the penalties are severe – termination of all Title XX funds to any state using those funds to purchase care in centers not meeting the FIDCR, and repayment by states of previously used Title XX funds. Clearly, regulations bearing this kind of penalty can cover only those characteristics of centers that policymakers deem absolutely essential to the welfare of children. In contrast, guidelines cover a range of center characteristics deemed desirable but not essential by policymakers, and do not carry penalties for noncompliance. Regulations require precise working, whereas guidelines can be worded broadly, with ample provision for local interpretation. By exploring different ways of defining and measuring center characteristics, the study will contribute to the content and wording not only of regulations but also of monitoring guidelines, purchase of service agreements, guidelines for center operations, and manuals for training and technical assistance.

Finally, it should be noted that NDCS research interests, of necessity, differ somewhat from the government's regulatory concerns, although common measurement needs have been stressed throughout this section. The researcher must be willing to expend the funds and effort to gather data both with a degree of detail and in a quantitative and qualitative form not needed by the regulator. Thus the data on center characteristics gathered during Phase II far exceed the requirements of most monitoring agencies, but they can be used to arrive at simpler and more policy-relevant descriptions of centers for future regulatory use.

STAFF QUALIFICATIONS

The first and most crucial "variable" in the child's day care experience is the caregiver.* Yet in many ways the characteristics of the caregiver are the most difficult and complex aspects of day care to define and measure. Warmth, maturity, concern, energy, humor, sensitivity, attitudes, motives and many other personal characteristics are important for work with children of any age and may be critical to successful relationships with three- and four-year olds. Center directors routinely make intuitive judgments about personal characteristics when they hire, train, supervise and evaluate caregivers. Yet it is difficult even to describe individual traits with scientific precision, much less write regulations specifying what configurations of traits are appropriate for work with children.

Regulatory agencies typically require qualifications such as formal education, previous experience and job-related training in an attempt to influence caregiver character and competence, however indirectly and imperfectly. The most important exception to this generalization is the recent adoption by several states of a program of caregiver certification based on systematic evaluation of competence on the job. This program, developed by the Child Development Associates Consortium served as the basis for one of the instruments used in Phase III data collection (described in Chapter 5).

**Ideally, the term "caregiver" should refer only to classroom staff – those who interact with children most frequently. In Tables 2-1, 2-2 and 2-3, however, "caregivers" also include center directors as well as a handful (47) of specialists and support staff. A separate analysis has shown little difference in the configuration of these Tables when only classroom staff are included.*

At the present time, many state codes are far more specific than the FIDCR in setting minimum standards of education, experience and training by job title. The FIDCR cite "training" as a crucial element in the background of the caregiver, but they do not distinguish formal education from other types of training, such as workshops and supervised internships, and do not specify what levels of education or types of training are desirable. Similarly, the regulations cite "experience" as an important background factor without distinguishing types or amounts of experience. Finally, the regulations do not specify different levels of training and experience for aides, lead caregivers and center directors.

Among the measurement issues raised by the FIDCR are: (1) What are the different types of education, experience and training that prevail among caregivers and must be considered in regulations and research? (2) What scales or categories are most useful for describing each type? (3) Can data on the different types be combined to yield meaningful overall indices of education, experience and training?

During Phase II, detailed information on most aspects of education, experience and training was collected via interviews with 672 caregivers from target classrooms. The data were used to compile a qualitative profile of caregiver backgrounds and to address the measurement issues just raised.

Education

Information was collected on total years of education, diplomas and degrees obtained, areas of specialization and education currently in progress. As shown in Table 2-1, one-quarter of the caregivers had at least a college education, while one-sixth had not completed high school. Almost half of those who had not completed high school were in the process of doing so, however. Of those caregivers who had some form of specialized training (in general only those who had some education beyond high school), more than half had concentrated in an area related to children. Thus, about 30 percent of all caregivers had specialized training in child-related areas. Similarly, more than 60 percent of those currently in some form of educational program were studying in child-related fields. About a third of all staff were currently working toward a diploma or degree.

Experience

Four different indices of experience were computed: (1) years of experience at the center where the caregiver was currently employed; (2) years of experience in current job (as distinct from experience in current center, where the caregiver may have held several jobs); (3) years of previous work experience as a caregiver; and (4) years of previous work experience with preschool children outside day care. Previous preschool experiences included Head Start, nursery school or kindergarten teaching, but did not include babysitting, church or summer camp experience. Years of experience in categories (1), (3) and (4) were adjusted to account for part-time jobs and jobs held simultaneously within the same category, and the adjusted values were combined to yield total years of child-related experience for each caregiver.

Table 2-2 shows the distributions of each measured experience variable among the 672 caregivers interviewed. About 56 percent had no previous day care or other group preschool experience before coming to work at the study centers. Half had less than two years of total experience, including their present job. Thus the Phase II centers were not characterized by staff with extensive previous experience, although some Atlanta centers had staff with considerably more experience than the norm.

Job-Related Training

Because of the number and variety of training programs available in the three study sites, it was not possible to compute a single meaningful variable representing the overall amount of training a staff member had.

Table 2-1

A. Education of Phase II Caregivers (N = 672)

Years of Education	Number of Caregivers	Percent of all Caregivers
0 - 11	107	16
12 - 15	394	59
16+	165	25
Not Reported	6	< 1

Degrees or Diplomas*	Number of Caregivers	Percent of all Caregivers
High School/GED	560	83
Voc./Tech./A.A.	89	13
B.A.	165	25
M.A./Ph.D.	30	4

* Note that some caregivers have more than one degree.

Area of Specialization	Number of Caregivers	Percent of all Caregivers
Day Care/Early Childhood Ed.	77	11
Elementary/Special Ed.	73	11
Child Development/Psychology	48	7
Other	172	26
None	302	45

B. Current Educational Activities of Phase II Caregivers (N = 236)

Expected Degree	Number of Caregivers	Percent of all Caregivers
High School/GED	47	7
Voc./Tech./A.A.	63	9
B.A.	85	13
M.A./Ph.D.	41	6
Total currently seeking degree or diploma	236	35

Current or Expected Area of Specialization	Number of Caregivers	Percent of all Caregivers
Day Care/Early Childhood Ed.	74	11
Elementary/Special Ed.	29	4
Child Development/Psychology	42	6
Other	77	12
Not Known	14	2

Instead, the data indicated the separate areas in which caregivers have received at least some training and areas in which they would like additional training. All but one of the ten specific training areas can be grouped under two broad headings – Child Development and Family (including day care teaching, child development/psychology, family life/counseling, human relations, and health/nutrition) and Curriculum and Activities (including use of materials in day care, day care curriculum, creative arts and child play/recreation). The tenth area is day care administration. As shown in Table 2-3, two-thirds of the staff had at least one course or other training experience in each of the

Table 2-2
Child-Related Experience of the Phase II Caregivers* (N = 672)

	Number of Caregivers	Percent of Caregivers Responding
Total Years of Experience		
Less than 2 years	302	49
2 - 4 years	135	22
More than 4 years	176	28
Experience at Present Center		
Less than 2 years	413	63
2 - 4 years	141	21
More than 4 years	106	16
Experience in Current Job		
Less than 2 years	459	69
2 - 4 years	127	19
More than 4 years	75	11
Previous Day Care Experience		
Some	189	29
None	459	71
Previous Preschool Experience		
Some	137	21
None	505	79
Day Care or Preschool Experience		
Either or Both	274	44
Neither	350	56

* Because of missing data, numbers of caregivers shown total less than 672; because of rounding, percentages may not total 100%.

Table 2-3
Training Achieved and Desired by Phase II Caregivers (N = 672)

	Number of Caregivers	Percent of Caregivers Responding
Training in Child Development and Family-Related Areas		
At least one course/experience	449	67
None	221	33
Training in Day Care Curriculum and Activities		
At least one course/experience	409	61
None	263	39
Desire for Additional Training		
Yes	558[a]	87
No	81[a]	13
Area in Which Training was Desired		
Child Development and Family	812[b]	61[b]
Curriculum and Activities	518[b]	39[b]
Field Placements		
Yes	280[a]	44
No	352[a]	56

[a] Because of missing data, numbers do not total 672.

[b] Each respondent was allowed to name three or more specific areas; hence responses total more than 672. Percents are based on all responses (1340).

two major areas and two out of five had participated in formally supervised on-the-job training. Eighty-seven percent of caregivers expressed a desire for additional training, primarily in Child Development and Family.

Comparison of Lead Caregivers, Aides and Directors

Not surprisingly, the qualifications of lead caregivers and aides differed significantly. Lead caregivers had more formal education (14.0 vs. 12.4 years), longer tenure in Phase II centers (2.9 vs. 1.4 years), longer tenure in their current jobs (2.3 vs. 1.2 years), and more previous child care experience (1.7 vs. 0.6 years). Directors averaged 16 years of education with the range of education extremely broad – from six to 20 years. Directors averaged 6.5 years of experience, and again the range was wide – from five months to more than 33 years. Only four directors, all in Atlanta, had no education beyond high school. However, more Atlanta directors (45%) had graduate training than did directors in Detroit (31%) or Seattle (19%).

Implications

As noted earlier in this chapter, the provisions of FIDCR relating to staff qualifications are ambiguous. While the FIDCR stipulate that caregivers "must have had training or demonstrated ability in working with children," specific educational requirements, levels of experience and types of training are not identified. The results of Phase II efforts to define and measure education, previous experience and in-service training of caregivers should be useful for future efforts to revise FIDCR provisions relating to staff qualifications.

Phase II data show that many day care center classroom staff have limited child care-related education and previous day care experience. Of the caregivers in the Phase II study centers, 56 percent had no previous experience working with children (aside from babysitting and other informal care arrangements) when they first came to work at their current centers. In addition, 71 percent had no formal education in child-related areas. While national data from the Supply Study will provide a firmer base for generalizations, it is clear from Phase II data that strict federal regulations pertaining to education and experience would severely affect day care centers' hiring practices. Many of the caregivers being hired under the current system would not meet requirements for extensive previous experience and/or formal child care-related education.

Phase II experience indicates that formal education is much more easily defined and measured than previous experience and in-service training. Collection of data on formal education and computation of an overall index of educational attainment are relatively straightforward. Formal education would therefore be easy to use as a regulatory standard. Years of education, degrees obtained and degrees in progress will be used as independent variables in Phase III analyses of the correlates of classroom process and child outcomes.

Experience presents somewhat more complex computational problems because different types of experience overlap in time, and the intensity of experience over a given period of time varies considerably. These facts necessitated use of rough estimation procedures during Phase II and some dependence on the caregiver's memory, potentially introducing error into the data. Nevertheless it was possible to compute a rough index of total years of experience that gave an approximate representation of each caregiver's background on this dimension. This index was provisionally adopted as an independent variable for Phase III, but exploratory work continues on the proper scaling of the experience variable. A variety of techniques will be examined in Phase III to test for differential effects of different amounts and types of previous experience. These will include tests for thresholds beyond which additional experience may influence child outcomes not at all or with reduced impact.

Training proved to be a particularly difficult variable to define and measure precisely. Training courses vary widely in content, format, duration, intensity and quality. Even with detailed probing, it was not possible to elicit training information in a form that permits comparison across different programs or that permits construction of a summary index. An obvious but provisional regulatory implication is that it will be difficult to write training standards that work well in varied environments. However, data collection procedures for Phase III have been refined and extended in a further attempt to standardize training information.

STAFF/CHILD RATIO AND GROUP SIZE

Staff/child ratio and group size are tightly interwoven. They are covered in the same provision of the FIDCR[5] and raise highly similar measurement issues.

The 1968 FIDCR, as modified by Title XX, mandate the following pattern of ratios and group sizes by age:

Age of Child	Minimum Staff/Child Ratio*	Maximum Group Size
Birth-6 weeks	1:1	Not Specified
6 weeks-3 years	1:4	Not Specified
3-4 years	1:5	15
4-6 years	1:7	20
6-10 years	1:15	25
10-14 years	1:20	25

Few states have staff/child ratio requirements as high** as those in the FIDCR. Most states permit ratios of 1:10 or less for three- through five-year-olds, in contrast to the FIDCR's 1:5 for three-year-olds and 1:7 for four- and five-year-olds. For children between the ages of three and five years, only three states have staff/child ratio requirements which match or exceed the FIDCR. In the case of six- to ten-year-olds, however, regulations in thirteen states match FIDCR, and regulations in nine more states exceed FIDCR.[6] Clearly the need for research on the appropriateness of the FIDCR ratios is greatest for children of preschool age; it is for that age range that federal and state requirements differ the most.

As indicated earlier, the rationale for the strict ratio requirement of the FIDCR is that high ratios permit caregivers to devote more time, energy and attention to each child than do low ratios. However, the rationale for the group size requirement is less clear. For any fixed level of staff/child ratio, variations in group size imply classrooms that differ in their organizational properties, possibly with different results for children. Ratios of 1:8, 2:16 and 3:24 are all the same numerically, but the dynamics of the three classrooms are likely to vary widely. In larger groups each child must form relationships with more peers, and these relationships may therefore be less intense. Similarly, since each caregiver is likely to feel responsibility for the group as a whole, relationships between the caregiver and the individual child are likely to be less intense than in the smaller group situation. Some children, espeically the shy or withdrawn, may escape notice and receive very little of the caregiver's time and effort, while others receive the lion's share of attention. This appears to be the reasoning underlying the FIDCR group size ceilings.

On the other hand, the opportunity provided by large groups for interactions with a larger number of children and, generally, a larger number of adults may be valuable. Larger groups usually have multiple caregivers. (Well over half of the Phase II classrooms had two or three caregivers working together, while only a quarter had just one adult.) Multiple-caregiver classrooms allow adults to share conversation, ideas, plans and responsibilities; moreover, they make it possible for one caregiver to take a break or attend to the needs of individual children while others continue to supervise the group as a whole. Thus problem children, both those who are disruptive and those who are withdrawn, may receive *more* individualized attention under a team-teaching arrangement. These alternative hypotheses about the effects of group size will be explored with the aid of classroom observation data on child-child and child-teacher interactions in Phase III.

In addition, ratio and group size may interact in complex and subtle ways. The ideal ratio for a small group may not be ideal for a large group. Such interactions will also be explored in Phase III to the extent that the data permit.

Though the FIDCR provision of ratio and group size at first appears to be quite specific,

---

**"Minimum" staff/child ratios specify the smallest number of adults who may be present with a given number of children. The Title XX FIDCR differ from the 1968 FIDCR in that the latter set no standards for children under three. Also, the 1968 FIDCR set one ratio (1:10) and one group size (25) for children from six through 14 years.*

***As noted in Chapter 1, "high" staff/child ratios imply fewer children per staff member, or a relatively large amount of staff time per child.*

it in fact leaves unanswered a number of questions regarding measurement. These include:

1) Who should be included in ratio and group size measurements — should part-time staff and children, "drop in" children, volunteer staff, or non-classroom staff be counted?

2) How should ratio and group size be assessed — by examination of center records or by direct observation?

3) Should record-based assessments be made prospectively (on the basis of schedules or rosters) or retrospectively (on the basis of attendance)?

4) On what part of the day, week or year should observed or recorded ratio and group size be based, assuming that ratio and group size fluctuate over time?

5) How should average ratios and group sizes be computed, again assuming some fluctuation?

In the subsections that follow, Phase II data are brought to bear in answering these questions. Throughout the discussion, choices are guided by the basic assumption that the reason for regulating ratio and group size is to enhance the quality of daily interactions between caregivers and children. While the most valid measure is always the one that appears to reflect the child's experience most directly, validity must often be weighed against the difficulty of collecting the needed data.

Who Should Be Included in Ratio and Group Size Computations?

Centers vary in stability of staffing and enrollment. In some centers, the same caregivers are present with the same children all day, all week, all year. In other centers, many staff and children are part-time, and come and go over the course of the day and week. Some centers experience rapid turnover of staff and children. Some serve "drop-in" children (children brought to the center sporadically, for brief periods, such as when the parent goes shopping or has a doctor's appointment), while others do not. Some centers, such as parent cooperatives, have many volunteer staff, while others do not. Almost all centers have staff who do most of their work outside the classroom — administrators, cooks, janitors — but many of these also double as part-time classroom staff. These variations in number of staff and children present must be addressed and dealt with uniformly in any revision of federal regulations.

In its analyses of ratio, the NDCS has generally counted all adults and children actually present in the classroom. The rationale for this approach is that all adults and children contribute to the dynamics of the classroom. For Phase II target classrooms, estimates of staff/child ratios based on classroom observations were not much affected by the inclusion or exclusion of specialists, volunteers and part-time children in the calculations. These target classes had been selected, however, because they included a large number of full-time children. In other day care centers, with a higher proportion of "drop-in" children, for example, estimated ratios would be more sensitive to the inclusion/exclusion of part-time children and unpaid or special staff. Phase III findings are needed before any final recommendations are made regarding which staff and which children should be included in ratio calculations.

Should Ratio and Group Size Be Assessed by Records or Observations?

There are clear advantages and disadvantages to choosing either center records or direct observations as data sources for computing ratio and group size. Observations are prima facie valid; they give a direct picture of the classroom social environment and if used for regulatory purposes in the form of random, unannounced spot-checks, they are difficult to falsify. They automatically take account of absentees. The main disadvantages of observation techniques are the intrusiveness that comes from the presence of observers in day care classrooms and the multiple visits that must be made to centers to obtain stable and reliable observation-based estimates of "typical" ratios.

Center records can provide either very detailed or very superficial measures of staff/child ratios and group sizes. If very detailed records are maintained by centers (the kind of records used for NDCS research purposes, for example), accurate information can be obtained on staff turnover, variations in grouping patterns across the day and week, the continuity of specific caregivers working with specific children, and absence rates for staff and children. Less detailed records may omit potentially important information — changes in staff/child ratios over the course of the day or week, absence rates, the degree of congruence between schedules for part-time staff and part-time children, or even within-center, classroom-specific staff/child ratios. It should be noted that, for monitoring purposes, reliance on center records rather than on direct observation shifts part of the burden for collecting monitoring information from monitoring agencies to day care center staff. In addition, use of center records requires monitoring agencies to train center staff in appropriate recordkeeping practices. The most obvious disadvantage of heavy reliance on center records is that it is difficult for monitors to verify the accuracy of recorded information.

In light of the complementary advantages and disadvantages of observed and record-based ratio and group size data, it may well be necessary for monitoring agencies to use both types of information in assessing compliance. It has been imperative in the NDCS, for example, to collect observation and record-based information in parallel throughout Phase III. The policymaker or monitor needs to know, of course, whether the two methods yield comparable information, and if not, what the different meanings of observed and record-based data are. This issue was explored in Phase II, and will be further explored in Phase III. Phase II results are discussed in the following subsections.

Comparison of Observed and Scheduled Staff/Child Ratios

Throughout Phase II, study centers prepared detailed standardized records of the number of hours per week children and staff in target classes were expected to be at centers. The ratio of weekly hours for staff to weekly hours for children provided a schedule-based measure of staff/child ratio. During the same period of time, classroom observations were used to generate an alternate measure of staff/child ratio. A series of observations was made for each target class at different times during the day and on different days of the week. For each observation, the number of adults present was divided by the number of children present to calculate an observed ratio for a specific sample period.*

In general, observation-based measures of staff/child ratios were higher (more staff for a fixed number of children) than schedule-based measures. The difference between the two measures was due to differences between scheduled numbers of children and staff in target classes and the actual numbers of children and staff observed to be present at various sample points. Observed ratios would be higher than scheduled ratios when absence rates for children (12% on average) were higher than absence rates for staff (negligible, on average) or when unscheduled staff were observed in target classes in higher proportion than unscheduled children.

For schedule-based ratios, hours were counted for all children and staff scheduled to be included in target classrooms for at least one hour per week. Volunteer staff hours were counted only if the volunteer worked in a target class for at least ten hours per week. For observation-based ratios, all children present at the time of the observation were counted. Adults were excluded from the calculations only if they classified themselves as bystanders (e.g., parents and others visiting the center, or non-scheduled staff stopping by momentarily).

Table 2-4 presents a comparison of observed and scheduled staff/child ratios for each of four months during Phase II. The number of children per caregiver is generally higher for scheduled ratios than for observed ratios, but the two measures converged over the course of the year. Because changes were made in data collection techniques during Phase II to increase the amount of information collected and to improve the quality of the data, the numbers in Table 2-4 provide only rough indications of the differences between schedule- and observation-based staff/child ratios and of the changes which are likely to occur over the course of a year. When further analyses have been conducted on data from Phase III and from the Supply Study, it will be possible to determine whether the patterns exhibited in Phase II data are representative of scheduled and observed ratios nationwide.

Table 2-4

Comparison of Observed and Scheduled Staff/Child Ratios

Month		Average Observed Ratio	Average Scheduled Ratio
October	1975	1:6.0	1:7.6
December	1975	1:6.8	1:7.3
March	1976	1:6.8	1:7.3
May	1976	1:7.3	1:7.1

Note: Observed ratios are based on classroom observations conducted during the *morning* hours on four sample days in each sample month. Scheduled ratios reflect the total staff and child hours scheduled for target classes during *all* hours of a sample week in each sample month.

It should be noted that schedule-based and observation-based ratios were highly correlated for Phase II target classrooms, despite the fact that scheduled ratios were generally lower than observed ratios. Classes with high scheduled ratios tended to have high observed ratios, and those with low scheduled ratios tended to have low observed ratios. The correlation coefficients between the two ratio measures ranged from 0.6 to 0.7 over the four sample months in Phase II. Even with a correlation as high as 0.7, however, only 50 percent of the variation in one measure can be predicted from data on the other measure. Because the two ratio measures are not more highly correlated, Phase III analyses may well be able to determine which of the two measures is most highly related to indicators of child outcomes.*

Comparison of Observed and Scheduled Group Sizes: Scheduled group sizes simply reflect the number of children slated to be in a particular room at a particular time. Observed group sizes reflect actual clusters of children supervised by one or more caregivers in some common activity. Thus, for example, the observer of Ms. Smith's group would count all the children actually under her supervision, even if some were formally assigned to another teacher, and would not count children who had been assigned to her if they were elsewhere.

Empirical and theoretical differences between the two measures of group size can be quite large -- especially in centers organized on an open classroom model. In one such center in Atlanta, for example, 88 children were assigned to a single large classroom. It was impossible to determine in advance whether and how these children would be organized into smaller groups under the direction of a single caregiver. In effect, the only definition of group size possible on an a priori basis was 88 children. During classroom observations, however, children were found to be clustered together in more tradi-

---

**If the two measures were more highly correlated, the power of statistical tests of which measure is most related to child outcomes would be very low. Even a correlation of 0.7 significantly reduces statistical power.*

tional, smaller size groups (average of 25.2 children per group). While groups separated and reformed over the day and week, the average size of the individual clusters remained quite stable.

An important question to be addressed by Phase III analysis, therefore, is which of the two measures of group size is most highly related to indicators of child outcomes. Scheduled group sizes tend to be higher than observed group sizes. In December, for example, average scheduled group size was 24.6, compared to an observed morning average of 17.3. Scheduled group sizes also varied more across centers than observed group sizes. Standard deviations – measures of variability – were 19.3 for schedule-based data and 5.4 for observation-based data. Since the correlation between observed and scheduled group sizes are quite low for the study centers, the power of Phase III analyses to differentiate the effects of the alternate measures will be quite high.

Should Measures of Ratio and Group Size Be Prospective or Retrospective?

Observed ratios and group sizes necessarily reflect the immediate situation in the classroom, but ratios and group sizes based on center records can be computed before the fact or after. Before-the-fact or prospective records – that is, rosters or staff and children scheduled to be present during each hour of the day – reflect which arrangements of staff and children centers consider best and the quality of such records may indicate the care with which centers plan. As shown above, Phase II data based on prospective records correlated well with what was actually observed in the classroom. However, the number of children observed was systematically overestimated because absentees were ignored.

Retrospective records which keep track of absentees, unexpected drop-ins and volunteers can describe the configuration of staff and children with complete accuracy only if they are kept meticulously and updated frequently – at least several times a day if they are to reflect the fluctuations of staffing and enrollment across major activity periods described below. The principal disadvantage of retrospective records is the bookkeeping effort required. Because this effort probably precludes widespread adoption of retrospective records, the NDCS has continued to use prospective records during Phase III.

When Should Ratio and Group Size Be Measured?

The question of when to measure ratio and group size arises only if these center characteristics vary systematically over the day, week or year. If they do not, measures will yield essentially the same results regardless of when they are taken, within the limits of random variation. If ratio and group size do vary systematically, a decision must be made as to which times of day, week or year are most important, or, alternatively, how to average across different time periods. The issue of systematic variation was examined extensively in Phase II.

Variation Across Segments of the Day: Perhaps the result that emerged most clearly from Phase II investigations of ratio and group size was that both variables are keyed to the activity pattern of the center and consequently fluctuate over the course of a day as activities shift. Table 2-5 illustrates the parallel fluctuations of activity, ratio and group size. The day begins with a period of free play while children arrive. (Some centers serve breakfast during this period.) Groups are small, since many children have not yet arrived, and ratios are high, even though some caregivers may also arrive later. The morning is split between planned group activities and free play. Group size stabilizes at about 17 and ratio at 1:7 on the average. During mealtime, beginning at noon or a little before, some centers bring several groups of children together while others do not, resulting in an average increase in group size across centers, and an increase in center-to-center variability. Naptime follows in the early afternoon; staff tend to leave or take breaks, so that ratio drops. Free play dominates the afternoon hours, though there is some planned group activity. During this time, group sizes and ratios return to approximately their morning levels, although in many centers there is a change in staff, often with paraprofessionals or students replacing lead teachers in the late afternoon.

This clear pattern is an essential feature of the day care environment, and any regulation involving ratio or group size must take it into account. Regulations must specify when and how often during the day ratio and group size are to be measured, and, if several time segments are to be used, how ratios are to be averaged. These requirements apply whether ratio and group size are assessed by observation or schedule. Fortunately, from the viewpoint of regulatory simplicity, there appear to be no important differences in ratio or group size across days of the week.[7] If this finding is confirmed in Phase III, selecting days or averaging across days need not be considered.

Stability and Change in Ratio and Group Size Over the Year: Stability of ratio and group size is an important issue for monitoring purposes. Particularly when ratio and group size are assessed by observation, both variables must be measured often enough that a fair and accurate general characterization of a center emerges, not a characterization distorted by unusual, nonrecurring events. A one-time observation might place a center out of compliance with the FIDCR if the observation happened to take place when several caregivers were sick, or when an unusual number of drop-in children were present. The stability of both observation-based and schedule-based measures of ratio and group size will be examined in great

Table 2-5

Group Activities, Group Size and Ratio*

Segment of the Day	Most Frequently Observed Activity	Average Group Size	Average Staff/Child Ratio
7:00 - 8:59	Free Play	11.6	1:4
9:00 - 11:29	Planned Group Activities	17.2	1:7
11:30 - 12:29	Meal	21.3	1:7
12:30 - 2:29	Nap	17.3	1:9
2:30 - 3:59	Free Play	17.1	1:7
4:00 - 4:59	Free Play	14.7	1:6

* Data for this table were taken from observations in target classrooms conducted during November and December 1975. The activities coded included arrival, free play (indoor and outdoor have been combined), planned group activity, individual activity, snack, meal, nap, transition from one activity to another, departure and other.

detail during Phase III. An important component of this analysis will be a determination of the relationship between the number of sample points at which data are collected and the potential sampling error in estimates of mean ratios and group sizes. This relationship has important bearing on the number of monitoring visits which are necessary to ensure reliable estimation of the degree to which an individual center is in compliance with federal and state staffing and group-size requirements.

Staff/child ratios declined slightly during the Phase II year; for example, average morning ratios fell from 1:6.0 in October 1975 to 1:7.3 in May 1976. At the same time, group sizes rose slightly; morning group sizes rose from 16.7 in October to 18.6 in May. Average number of caregivers per group was 2.4 in both cases. This modest trend is apparently due to a general tendency for centers to expand enrollment by small numbers of new children as the year progresses.

Despite the modest seasonal trend, centers with high ratios in one month remained high in other months, and centers which were low in one month remained low. Similarly within any one observational period, centers that had high ratios at any one time of day also tended to have high ratios during other segments of the day. Group size was also consistently high or low across periods of the day and months of the year.

During Phase III the NDCS will attempt to confirm the results by once again examining ratios and group sizes systematically, both by observation and by consulting center records, at four time points during the day (morning, lunch, nap and afternoon) and at five time points evenly spaced over the course of the year. The results presented here, of course, do not reflect deviations which may occur during summer months. In addition, the results pertain to centers included in the Phase II study sample and to target classes with high proportions of full-day, full-week children. When Supply Study data have been analyzed – especially classroom observation data obtained for a national probability sample of 70 centers – a more adequate basis will be available for generalizations about the stability of alternate measures of ratio and group size across different types of centers.

How Should Average Ratios Be Computed?

Once the issues of how, whom and when to count have been settled, the regulator must decide how ratios and group sizes should be computed. No problem arises for group size, but computation of ratios is not as straightforward as it may seem.

The form of ratio computation that is simplest and imposes the fewest demands in terms of data collection is a ratio of total number of staff to total number of children enrolled. However, this crude head-count ratio ignores variations in both the number of hours that children and staff are scheduled to be present in the classroom and whether they are actually present. A higher level of precision is obtained with ratios based on scheduled or actual hours of attendance, but this requires additional decisions regarding methods of averaging.

If one considers only a single-ratio measurement, computation is deceptively simple; one counts children and staff and divides one number by the other. It matters little which number one picks as numerator and which as denominator – that is, whether one computes staff/child or child/staff ratio – since the two are merely inverses of one another. However, the fluctuation of ratio across activity periods of the day confronts the regulator with the need for a formula to average disparate ratios that exist for different lengths of time. Simplicity evaporates when different averaging techniques are considered. Child/staff and staff/child ratios do not yield the same results when averaged. For example, the average child/staff ratio for May classroom observations in the 64 study centers is 8.3:1, compared to the average staff/child ratio shown in Table 2-4 of 1:7.3. If a monitoring agent were to average ratios for different times of the day, for different days of the week or for different classes in a particular center, the results could be quite different depending on whether child/staff or staff/child ratios were used in the computation.*

The complexities of alternative approaches to computing staff/child ratios are best presented in terms of a specific example. Consider a hypothetical day care center classroom with 17 children enrolled – ten children for eight hours each weekday, two children for four hours on Monday, Wednesday and Friday mornings, and five afterschool children for four hours each weekday afternoon. Care is provided by four adults – a full-time caregiver working eight hours each week day, a part-time person working four hours on Monday, Wednesday and Friday morning, a part-time person working when afterschool children are present, and a volunteer working eight hours on Wednesday. The adult/child schedule for the center would appear as follows:

Adults/Children

	Mon	Tues	Wed	Thur	Fri
AM	2/12	1/10	3/12	1/10	2/12
PM	2/15	2/15	3/15	2/15	2/15

A variety of methods can be used to compute staff/child ratio for this classroom. Some of the most obvious and most often used methods are the following:

- *Head Count Ratio:* Ratio of the total number of adults providing care to the total number of children enrolled. In this example, the computed ratio would be 1:4.3, if the volunteer is included, or 1:5.6, if he/she is excluded.
- *Hour-on-Hour Ratio:* Ratio of the total number of adult hours to the total number of child hours; in this example, the hour-on-hour ratio is 80/524 or 1:6.6. If volunteer hours are excluded, the estimated ratio would be 1:7.3.
- *Average Staff/Child Ratio:* Another ratio estimate can be calculated by computing the ratios prevailing for each of the five AM and each of five PM periods and then computing the simple arithmetic mean of the ten numbers. The resulting average is 1:6.6, or 1:6.8 if the volunteer is omitted from the calculations.
- *Average Child/Staff Ratio:* Instead of averaging staff/child ratios across the ten AM/PM periods, one can average child/staff ratios. A different ratio estimate is obtained – 7.1:1 with the volunteer or 7.3:1 without the volunteer.
- *Observed Staff/Child Ratio:* Classroom observations at specific times during the week can be used to estimate ratios. If no staff or children are absent at the time of the observation, one would observe a 1:10 ratio on Tuesday and Thursday mornings, a 1:6 ratio on Monday and Friday mornings, a 1:4 ratio Wednesday morning, a 1:5 ratio Wednesday afternoon, and a 1:7.5 ratio in the other afternoons. If children and/or staff are absent, one might find ratios between 1:3 on Wednesday morning and 1:15 some afternoons. Since children are more likely to be absent than staff, observed ratios should be higher (fewer children per caregiver) than scheduled ratios in most instances.

Averaging usually involves adding up the values obtained and then dividing by the number of values (arithmetic mean). There is an alternate averaging technique, seldom used and more difficult to compute, which produces the same result for child/staff and staff/child ratios. This average, known as the geometric mean, is calculated by* **multiplying *all of the values obtained* $(RxR_2xR_3...)$ *and then taking the 1/nth power of the product (n is the number of values used to calculate the average). The geometric mean for both child/staff and staff/child ratios from May classroom observations is 7.8 children per caregiver.*

When these alternative computations were applied to data for the 64 Phase II study centers, average ratios across all centers ranged from 1:5.3 for the hour-on-hour method to 1:8.3 for the average child/staff ratio method, a difference of 55 percent. Different computation methods produced different results for different centers; a center with a relatively high hour-on-hour staff/child ratio, for example, might have a relatively low head-count staff/child ratio.

Several general conclusions emerged from Phase II experience with alternate computation schemes which should be useful to agencies responsible for monitoring compliance with staff/child ratio requirements. First, computations based on simple head counts of staff and children are very unreliable indicators of the amount of caregiver time available to the average child. The presence of large numbers of part-time staff or part-time children will cause head-count ratios to seriously distort staff/child contact time. Second, superficial adjustments to the head-count method (for example, grouping children into part-time and full-time categories and then counting a part-timer as half a full-timer) improve the correlation between estimated ratios and staff/child contact time for most centers but still produce considerable distortions for a few centers. Third, computations based on prospective schedules of time-in-center generally underestimate staff/child contact time because absence rates are higher for children than for caregivers. Finally, which computation method is chosen has a very significant impact on estimated ratios and, thereby, on estimates of the degree of compliance with federal and state regulations.

As already noted, the NDCS is using both observations and center records to compute ratios in Phase III. Record data will be used to compute hour-on-hour estimates of staff/child ratios, while observation data will be used to compute averages of staff/child ratios across the day, week and year. Both techniques will also be used to compute group size estimates. The final Phase III report will demonstrate the implications of choosing alternate measures of ratio and group size not only in terms of which measure best predicts child outcomes, but also in terms of the costs of achieving compliance under the alternate monitoring definitions and measurement techniques.

OTHER REGULATABLE CENTER CHARACTERISTICS

In addition to the major policy variables — staff/child ratio, group size and staff qualifications — the NDCS explored several other currently or potentially regulated characteristics of centers during Phase II. These characteristics included (1) the physical environment (space, equipment and materials); (2) center philosophy and program orientation; (3) presence of supplemental programs and specialists in education, nutrition, health, medical care and social services; (4) opportunities for parent involvement; and (5) continuity of care (defined later). The purpose of the Phase II explorations was to determine (a) whether these characteristics could be measured and monitored accurately; and (b) whether any appeared to be sufficiently important to replace one of the major policy variables in the Phase III design. In general, the Phase II analyses led to the conclusion that these characteristics merited further study in Phase III but should not replace ratio, group size and staff qualifications as primary independent variables around which Phase III would be designed.

The Physical Environment

While most developmental psychologists and child care practitioners agree that the human environment is a more important influence on the growing child than the physical environment, space, equipment and materials can facilitate or impede caregiver-child interaction. In the extreme case, a crowded or inadequately designed facility, or ill-chosen equipment and materials, can be threats to safety. While licensing requirements in most states protect children from grossly unsafe environments, the question remains whether federal regulations might further strengthen these safeguards or even influence development in a positive way.

The primary source of data on space, equipment and materials in Phase II was the Physical Environment Inventory, which is a part of a larger instrument designed by the Stanford Research Institute (SRI) to record classroom processes. This instrument is described in Chapter 3, and Phase II data on the physical environment are presented there. To anticipate that discussion briefly, most of the Phase II centers and classrooms had adequate-to-good space, equipment and materials available. However, inadequacies in one or more of those aspects of the physical environment were noted for a small minority of classrooms. In addition, the quality of the physical environment was higher in centers serving children from families of relatively high socioeconomic status than in centers serving children from less advantaged backgrounds.

The data suggest that it will be possible to construct a rough index of the overall quality of the physical environment for use in Phase III analyses, although it will not be possible to construct a meaningful index involving fine gradations of quality. A simple dichotomy of classrooms or centers with and without significant physical deficiencies may prove to be the most useful and reliable form of index. The study of the physical environment will be augmented by two new data sources in Phase III: SRI has developed an instrument to assess components of caregiver competence, as specified in the Child Development Associate (CDA) certification system mentioned earlier in this chapter. This instrument contains a number of ratings of specific aspects of the physical environment. In addition, caregivers will rate the importance and valence (positive or negative quality) of the environment as part of an instrument called the Day Care Forces Inventory (DCFI) which is described further in Chapter 5.

From a regulatory perspective, the Phase II data reveal no clearcut benchmarks of quality in the physical environment. Rather, there is a tendency for a number of physical properties of centers to vary together. Because the regulatory utility of a global quality index based on ratings of aspects of the environment is questionable, there appeared to be no justification for altering the Phase III design by replacing any of the three major policy variables with an environmental index. However, data on the physical environment will be examined in conjunction with ratio, group size and staff qualifications in an effort to describe any interactions that may exist between the physical and human environment.

Center Philosophy and Program Orientation

The philosophy of a center and its general approach to the care and education of young children can potentially have profound effects on the dynamics of a classroom, regardless of the qualifications of caregivers or the configuration of staff and children. While direct control of the philosophy or "curriculum" of day care centers has never been seen as an appropriate regulatory approach for the federal government, the 1968 FIDCR did require that day care programs have an identifiable educational component. This requirement was not included in the Title XX FIDCR. An important policy question at the present time is whether such a requirement is a meaningful and useful way to build a mandate into regulations that federally subsidized day care should actively foster children's development. Several Phase II findings bear on this question.

During Phase II, data on center philosophies and curricula were gathered by means of interviews with directors and caregivers. The data were subsequently confirmed by Abt Associates field staff who had extensive experience in all of the study centers. The program typology that emerged from this process was based on expert judgment and not on an explicit, mechanical classification procedure. Programs were first sorted into two mutually exclusive categories, depending on whether they emphasized the individual child or the group. Individual emphasis implied stress on independence, self-reliance, learning at the child's own pace and making the child feel important or special. Group emphasis implied stress on sharing, cooperation, getting along with others and interaction between staff and children. Twenty-eight of the 64 Phase II centers were marked by emphasis on the group, and 14 by emphasis on the individual. Twenty-two did not clearly fall into either category and were classified as mixed.

Beyond this broad distinction between group and individual orientations, centers were categorized according to other program emphases. The latter categories were not mutually exclusive; centers could be placed in several categories if their philosophies or curricula seemed to involve multiple emphases, and the vast majority of centers were in fact placed in two or more categories. These additional descriptive categories were: (1) *Ethnic and/or religious emphasis* – applied to centers that stressed spiritual development or the importance of learning to appreciate one's own culture; (2) *Guidance emphasis* – applied to centers where staff stressed love, warmth, security, understanding or a "homey" atmosphere; (3) *Protective emphasis* -- applied to centers that saw themselves principally caring for children's basic needs in order to free their parents for work or other activities; (4) *Behavioral supervision emphasis* – applied to centers that stressed obedience, discipline, manners and adult supervision and control; and (5) *Cognitive emphasis* – applied to centers where staff stressed such activities as language stimulation, or teaching ABC's and number concepts, and where the day care classroom was described as preparation for the grade school environment.

Two major conclusions emerged from the construction of this typology. First, day care personnel are concerned with preparing children for school, but they define school readiness broadly, including preparation for the social environment of the school as well as enhancement of academic skills. Of the 64 centers, 52 emphasized stimulation of cognitive skills, but only two stressed cognitive skills exclusively. The remaining 50 stressed cognitive skills in conjunction with one or more of the various approaches to social development listed above. Because cognitive and social emphases are both seen as part of preparation for school, it is difficult to separate the educational component of day care from center activities in general. However, it may prove useful in Phase III to examine differences between the few centers that did *not* include cognitive skills among their major emphases and the majority of centers that did.

Second, and more generally, it was not possible to construct a clearcut typology of mutually exclusive categories that fit all centers. On these pragmatic grounds alone, it would have been inappropriate to base the design of Phase III on a typology of program philosophies or curricula. Moreover, most of the differences in program emphasis that distinguished the Phase II centers are not relevant for regulation; for example, choosing between centers that emphasize groups or individuals, choosing a center that emphasizes ethnic identity, or choosing one that stresses strict discipline are matters of parental discretion, not governmental policy. Thus, from a regulatory perspective as well as a pragmatic one, examination of program differences deserves a secondary role in Phase III – with the possible exception of the difference between programs that do and do not have some cognitive emphasis, as discussed above.

Supplemental Services and Specialists

The FIDCR require that services from specialists and health and social agencies must be made available to children and families through the center. The FIDCR section on social services requires that:

- Social services be provided in the center or by an administering or operating agency;

- Counseling and guidance be provided to help the family determine appropriateness of day care;
- Assessment be made continuously of the child's adjustment to day care; and
- Procedures be established for coordinating services with other agencies.

In addition, federal requirements concerning health and nutrition also include provisions for:

- Ensuring general health and safety of the child and his environment;
- Arranging for medical health evaluations and care appropriate for the child's age;
- Arranging for immunizations appropriate for the child's age; and
- Serving adequate meals and snacks.

Directors of the Phase II centers were asked whether their centers provided or arranged for each of the following services: emergency medical care, other medical care (followup to previous treatment, treatment of minor illnesses or infections), immunizations, other preventive health services (health screening, tests of hearing and vision) and social services. Directors were also asked whether their centers employed specialists in developmental testing, education and nutrition. Emergency medical services were provided by all but five of the 64 centers. Other medical and social services were provided by three-quarters of the centers. Immunizations were available in 34 of the centers; other preventive health services were available in 30. Forty-eight of the centers used specialists in developmental testing. Forty-two had specialists in nutrition. Thirty of the centers employed educational specialists. Immunizations, preventive health services and social services were more prevalent in centers serving low-SES families and in centers funded principally by the government than in other centers. Educational specialists were also more prevalent in publicly funded centers.

Many of these supplemental services relate only indirectly to most of the child outcomes under investigation in the NDCS (described in Chapter 3), although availability of such services may affect parent satisfaction (also discussed in Chapter 3). Moreover, whether centers need to provide such services, even for low-income families, depends on what other resources are available in the community. For example, immunizations may not be necessary if they are readily available through a free clinic in the neighborhood. For both of these reasons, availability of specialists and supplementary services were not appropriate center characteristics to serve as primary independent variables in Phase III.

Opportunities for Parent Involvement

The FIDCR require that parents be given the chance to work cooperatively with staff and to observe their children in the center. They also stipulate that parents must have the opportunity to participate in making decisions about the center and its operation. Finally, centers are required to provide guidance and counseling services; for example, staff are to help families assess the appropriateness of their day care arrangements and plan medical and dental services for their children.

During Phase II, center directors were asked whether they provided opportunities for parents to confer with staff, to attend meetings of the center's governing board, to participate in educational groups, to participate in social activities and to work as volunteers in the center. All but six of the centers had parent-staff conferences. Twenty-nine allowed parents to attend board meetings. Forty-four provided parent education groups and fifty had organized social activities for parents. Thirty-nine provided opportunities for parents to work as volunteers. Except for parent conferences, which were nearly universal, all of these forms of parent participation were more prevalent in publicly funded centers than in parent-fee centers.

This information provided by directors reflects only the opportunities for parent involvement available in the centers. Whether parents are aware of such opportunities and take advantage of them is a separate question. Data on actual involvement from the parents' perspective are presented in Chapter 3.

Like many supplemental services, most forms of parent involvement cannot be expected to exert a direct influence on center processes and child outcomes, although they may be directly related to parent satisfaction. While parent volunteers may influence the behavior of caregivers and children directly, indirect influence is the most that can be expected from parental participation in decision-making, parent-staff conferences and parent education programs. Parent participation in social activities may have no influence at all. Consequently, it was not appropriate to substitute parent involvement for one of the major policy variables in the Phase III design.

Continuity of Care

Many developmental psychologists argue that the availability of a single caregiver over a long period of time is a more important feature of the growing child's human environment than the number of other children who share the caregiver's attention. While this argument is usually applied to infants and toddlers, it may apply to preschoolers as well. Thus, as a supplement to the basic analyses of ratio and group size, the NDCS is examining the effects of the number of different caregivers to whom the child is exposed over the course of the day, week or year. Such an analysis was not possible in Phase II because ratio observations recorded only the numbers of caregivers present, not their identities. When the possible importance of continuity of care was recognized, Phase III ratio observation procedures were modified to include names of caregivers. (This change in procedure will also allow the NDCS to examine fluctuations in the kinds of staff present during different periods of the day — for example, differences in education levels between morning and afternoon staff.)

Summary

The physical environment, center philosophy and curriculum, supplementary services, parent involvement and continuity of care may all influence the outcomes of day care for children and families. However, the physical environment and center philosophy and curriculum pose problems for measurement and regulation, while supplementary services and most aspects of parent involvement are likely to affect parent satisfaction more than center processes and child outcomes. Since the NDCS focuses on the behavior of children and caregivers in centers and on changes in children's cognitive and socioemotional development (as well as on parent satisfaction), none of these characteristics appeared to merit a central role in the design of Phase III. Continuity of care is so closely allied to ratio, group size and staff qualifications that it could be incorporated without major modification of the design.

Information on all of these variables will be integrated into an overall description of the operation of the Phase III centers. Relationships between these variables, the major policy variables, center processes, and child and family outcomes will be described. However, because Phase III is designed to highlight the effects of the policy variables, it will not be possible to perform definitive analyses of the effects of other regulatable characteristics. Some of these characteristics tend to occur together, in broad configurations, rather than to vary independently. Specifically, it is in publicly funded centers serving low-income people that many supplementary services are offered, many specialists employed, and many opportunities for parent involvement offered. Phase III will not allow a full separation of the effects of public funding, family SES, services, specialists and parent involvement. The relationships among parent involvement, supplementary services, public funding and family SES underscore the importance of a Phase III experimental design that clearly separates the effects of the policy variables, especially staff/child ratio, from all of the other factors with which the policy variables are usually associated in the day care world.

Chapter 3: PROCESSES AND OUTCOMES IN DAY CARE

Underlying the FIDCR are two basic assumptions — that interaction between child and caregiver affects the child's development, and that the quality of adult-child interaction varies with the human and physical organization of the day care center. As noted in Chapter 2, the implicit rationale for regulating staff/child ratio and group size is that high ratios and small groups permit caregivers to attend to the emotional and educational needs of individual children. Similarly, the rationale for regulating caregiver qualifications is that trained, experienced and/or educated caregivers are likely to recognize and meet children's needs, enhancing both children's development and parents' satisfaction.

In order to test these fundamental assumptions, the NDCS is examining the impact of currently and potentially regulated characteristics of centers and caregivers on: (1) the behavior of caregivers toward children; (2) the behavior of children in the day care setting; (3) other measurable aspects of children's socioemotional and cognitive development; and (4) parent satisfaction and parent involvement in center activities. This chapter discusses measurement of these center processes – behaviors of caregivers and children — and outcomes for children and their families.

The developmental literature available in 1974, when the study began, did not provide direct, definitive guidance as to which elements of caregiver behavior in the day care setting are critical for children's development, nor which aspects of development are affected by variations in the day care environment. Consequently, the testing contractor, Stanford Research Institute (SRI), constructed a list of outcome variables specifically tailored to the purposes of the NDCS. SRI based its list on a review of literature on early childhood intervention programs, nursery schools and parent-child interaction, and on consultation with experts in day care and child development. The selection of child variables is described in the first section of this chapter. The section also describes the selection of measurement instruments.

SRI chose direct naturalistic observation* as the main data source for describing the behavior of children and caregivers in centers. The observation systems selected by SRI to record behavior of caregivers and children in day care settings are described in the second and third sections of this chapter.

In addition to direct observation of children's behavior, the NDCS used several standardized tests and rating scales to measure aspects of socioemotional and cognitive development. Psychometric properties of the Phase II test battery and rating scales are discussed in the fourth section.

Finally, structured interviews were used during Phase II to assess parents' satisfaction with their day care arrangements and parents' involvement in different aspects of center operations. The parent interviews are discussed in the fifth section.

The variables, instrument battery and data collection procedures were modified for Phase III on the basis of Phase II results. The modifications are described in Chapter 5, which deals with Phase III design and instruments.

The present chapter includes descriptive data; however, as noted in Chapter 1, no information on the effects of the policy variables is included, since all such information is preliminary and must be confirmed in Phase III. Summary profiles of the NDCS population on each measure are presented here, along with breakdowns by age for all child measures and breakdowns by sex for observational measures of children. The

---

**In naturalistic observation, the observer records the everyday behavior of the child or caregiver in the day care setting, without any deliberate intervention. Naturalistic observation contrasts with structured observation in which the adults or children being observed are presented with a standard stimulus situation.*

reasons for presenting this information are (1) to further describe the caregivers, children and parents in the NDCS sample; (2) to describe the physical and social environment of the NDCS classrooms; (3) to show that the measures retained for Phase III had adequate reliability and variability in Phase II, and (4) by comparing NDCS results to those of previous studies, to confirm both the validity of the measures and their sensitivity to well-established causal factors (e.g., the effects of age and sex on various child measures).

SELECTION OF CHILD VARIABLES AND INSTRUMENTS

A major mission of the NDCS is to determine what happens to children in day care as a consequence of differences in regulatable center characteristics. "What happens" refers both to the experience of the child in day care (process) and to changes in the child as a consequence of that experience (outcomes).

Identification of child process and outcome variables was initially the responsibility of the testing contractor, SRI. During Phase I, SRI both consulted a panel of experts in child development and day care, and conducted a review of the literature on effects of group care environments on the preschool child. This was done in an effort to identify variables that were:

- Likely to be affected by regulatable center characteristics, as evidenced by empirical research or suggested by plausible theory;
- Important to one or more of the major audiences of the NDCS – policymakers, parents, advocacy groups and the research community;
- Likely to show appreciable change in preschool children over the duration of the study; and
- Likely to be measurable within the time and resource constraints of the study.

SRI's final list of variables was formulated with the assistance and approval of Abt Associates and the Office of Child Development. These variables are listed in the first column of Table 3-1. Each variable is defined and discussed in SRI's report on the field testing of instruments for Phase II.[1]

The original list of child variables was re-examined on the basis of preliminary analysis of fall 1975 data. The second column of Table 3-1 shows the revised list of variables that resulted from this re-examination. The revised list has been provisionally adopted for use in Phase III. However, as explained later in this chapter, the list may be revised further, since observation data permit considerable flexibility in construction of summary variables.

The revised list incorporates the following changes: Two of the observation-based variables, *social involvement* and *autonomy*, were regrouped or relabeled to reflect more accurately the meanings of the observed behaviors, resulting in creation of variables called *interest/participation* and *patterns of interaction* (defined in a later section).
A new observation-based variable, *apathy/withdrawal*, was added to the list. Several observation-based variables (indicated by footnote "c" in Table 3-1) were not analyzed in Phase II because the relevant observations were very rare. (Adjustments have been made in Phase III data collection procedures in an effort to increase the frequencies of these observations; hence the variables were provisionally retained for Phase III.) Several of the standardized tests and rating scales (indicated by footnote "e" in Table 3-1) proved inadequate for NDCS purposes; these instruments and the variables associated with them were dropped from further analysis in Phase II and dropped completely from Phase III. Finally, measures of vocabulary and motor skill development were added to the Phase III test battery. Each child variable is discussed more fully in the context of the descriptive results presented later in the chapter.

Table 3-1
Phase II Child Outcome Variables and Instruments

Variables		Instruments	
Original List[a]	**Revised List**[b]	**Phase II**	**Phase III**
Dependency	Dependency	Prescott-SRI Child-Focus Observation Instrument (CFI)	CFI
Aggression	Anger/Hostility	CFI Kohn-Rosman Problem Checklist	CFI
Compliance/Obedience	Cooperation/Compliance	CFI Kohn-Rosman Social Competence Scale	CFI
Social Involvement	Patterns of Interaction	CFI	CFI
Autonomy	Interest/Participation	CFI Schaefer Day Care Behavior Inventory ("Extraversion" subscale) Kohn-Rosman Social Competence Scale	CFI
———————	Apathy/Withdrawal	CFI Kohn-Rosman Social Competence Scale Kohn-Rosman Problem Checklist	CFI
Self-Control	Self-Control[c]	CFI Motor Inhibition Test	CFI (Structured Situation)
Prosocial Behavior	Prosocial Behavior[c]	CFI Schaefer Day Care Behavior Inventory ("Considerateness" subscale)	CFI (Structured Situation)
Social and Emotional Harm	Social and Emotional[d] Harm	CFI Kohn-Rosman Problem Checklist Center Records	CFI
Task Persistence	Task Persistence	CFI Schaefer Day Care Behavior Inventory	CFI

(continued)

Table 3-1 (continued)

Variables		Instruments	
Original List[a]	Revised List[b]	Phase II	Phase III
Reflectivity[e]	—	CFI Matching Familiar Figures Test	—
Generation of Ideas	Generation of Ideas[c]	CFI	CFI (Structured Situation)
Problem Solving	Problem Solving[c]	CFI	CFI (Structured Situation)
Curiosity	Curiosity[c]	CFI	CFI (Structured Situation)
Language Skills	Language Skills	Preschool Inventory (64-item)	Peabody Picture Vocabulary Test (Shipman-Tanaka Revision)
Memory Skills[e]	—	McCarthy Verbal Memory Scale	—
Reasoning Skills and Concept Mastery	Reasoning Skills and Concept Mastery	Preschool Inventory (64-item)	Preschool Inventory (32-item)
Self-Help Skills[e]	—	Vineland Social Maturity Scale	—
Safety and Health[f]	Safety and Health	Center Records	Center Records
—	Motor Skills	—	Motor Skill items from Denver Development Screening Test

[a] Variables in this column were on the list developed in Phase I by SRI.

(continued)

Table 3-1 (continued)

[b] Variables in this column reflect a revision of SRI's list after preliminary analysis of the fall 1975 data. This list will provisionally be used in Phase III, although exploration of potential variables continues.

[c] The indicated variables were not analyzed in Phase II because the relevant observations occurred with extremely low frequency. As discussed in Chapter 5, structured stimulus situations will be used during Phase III in an effort to increase frequencies of these observations.

[d] Social-emotional harm crosscuts other outcome variables. A pattern of high rates of *dependency, anger/hostility, apathy/withdrawal* or physical injury, as revealed by center accident records, indicates potential harm.

[e] The indicated variables were not analyzed in Phase II because the relevant measurement instruments proved to be unreliable or, in the case of the McCarthy Verbal Memory Scale, showed ceiling effects. These variables and instruments have been dropped from Phase III.

[f] *Safety and health* were not measured in children directly but were assessed by examination of center records concerning accidents and illnesses. In Phase III, ratings of the adequacy of health and safety facilities and procedures in centers have been added.

Instruments for measuring child processes and outcomes were selected by SRI using criteria of reliability, validity and demonstrated practicality for large-scale field studies. A candidate set of instruments was identified and screened according to these criteria on the basis of a literature review and consultation with experts. Those instruments that survived this screening were then subjected to a comparative field test, and a final set of instruments selected. Wherever possible, multiple measures were used for the same variables, in order to permit cross-validation of both the instruments and preliminary findings about effects. The list of Phase II instruments appears in the third column of Table 3-1. The final list of instruments to be used in Phase III appears in the last column of the table. Each Phase II instrument, and the reasons for its selection, is discussed in SRI's report, *Phase II Instruments for the National Day Care Study.*[2]

SRI selected two general types of instruments for most Phase II child data collection – a set of standardized tests and an observation system. In addition, several rating scales to be completed by caregivers were field-tested in Phase II for possible inclusion in Phase III. For all three types of instruments, the first concern in Phase II was to determine whether reliabilities and other psychometric properties were adequate when the instrument was used under NDCS field conditions. For those instruments that showed adequate reliabilities, a second concern was to discover whether scores varied with known determinants of the child characteristics being measured, such as age or sex. As noted above, convergence with well-established previous findings tended to confirm instrument validity. In deciding to retain an instrument for Phase III, evidence of reliability was essential and evidence of validity highly desirable.

In addition to general concerns of reliability and validity, the observation instruments raised a number of more specific analytic and interpretive issues. These are discussed in the following paragraphs in order to provide the reader with a framework for understanding subsequent sections in which data from the observation system are presented.

Behavior Codes and Summary Constructs: Observation systems record the activities of children and caregivers* in the form of a large number of codes, each corresponding to a particular behavior occurring at a particular moment. By focusing on concrete behavior description, such instruments can achieve a relatively high degree of inter-rater agreement (compared, for example, to more global rating scales), as well as provide a precise, detailed picture of activities over the course of the observation period. However, in order to address more general variables of the type selected as child outcomes for the NDCS, it is necessary to combine particular behavior codes into summary constructs. For example, specific codes denoting a child's helping, sharing, and offering of sympathy or comfort might be combined into a more general construct such as *prosocial behavior.* This means that whenever any one of the component behavior codes appeared on a record of the child's behavior, it would be scored and analyzed as an instance of *prosocial behavior.* Combining behavior codes into summary constructs offers considerable economy in analyzing data and makes it possible to use observation data to capture broader variables that have been assessed by other techniques in previous research. Moreover, new constructs can always be developed from the information recorded by an observation system. Thus, the NDCS has considerable flexibility to examine new outcome variables that may arise in the literature or be suggested by consultants. For this reason, exploration of new child outcome variables can and will be pursued during Phase III.

Rare but Important Codes: Certain behaviors of theoretical and practical significance accounted for only a very small proportion of the Phase II observations. Some were negative events which are potentially important as indicators of harm, such as extreme emotional distress, fighting and tantrums on the part of children, or physical mistreatment of children by caregivers. Others were positive events which are potential indicators of the benefits of day care, such as self-control and voluntary sharing and helping by children.

In some cases these behaviors were genuinely infrequent. However, in other cases these behaviors were seldom recorded, not because they were infrequent, but because they were of very short duration. When observations occur at regular intervals, as did the NDCS child observations, brief activities may be recorded rarely or missed altogether. Whatever the reasons that key behaviors were infrequently recorded, their low frequencies required special techniques to analyze the resulting data. (These are discussed in Appendix A.) During Phase III, structured stimulus situations will be used to increase the frequencies of rare but important codes. The structured situations are discussed in Chapter 5.

Traits and Situations: The behavior of children and adults in natural settings reflects both the enduring characteristics of the individuals involved and the opportunities and incentives provided by the situation. Thus, for example, when an angry child is observed, it is difficult to determine without additional information whether the child is characteristically belligerent or whether his/her anger is due to some frustration, threat or infringement of rights. For NDCS purposes, measurements of child traits and of situational pressures are of equal importance. The traits of children may change as a consequence of different types of care, and thus traits are outcome variables. The kinds of situations in which children find themselves – stimulating or boring, threatening or reassuring – are part of the day care process and may also be affected by such center characteristics as staff/child ratio, group size and staff qualifications.

Observer Effects: Observation data inevitably reflect characteristics of the observer to some

---

**In addition to the observation system used to describe the behavior of children, SRI selected a second system to describe the behavior of caregivers. The latter system is discussed in a later section.*

degree. Different observers have different perspectives and will record the same events at least somewhat differently. Moreover, observers may affect the behavior of the people they observe; thus, the events seen by different observers may actually be different in systematic ways. These observer effects are minimized when observers are carefully trained and codes are carefully defined, as was the case with SRI's training procedures and instruments. However, observer effects cannot be eliminated entirely and must be taken into account in assessing the meaning of observation data.

Preliminary Phase II analyses suggested that differences among observers assigned to different classrooms or centers may indeed contribute to measured differences in the behavior of caregivers and children. One such observer difference potentially important for future policy research was linked to the race of the observer; small but systematic and significant differences emerged in codes used by black and white observers, for both the child and adult observation instruments. These differences were too small to threaten the basic policy analysis. Moreover, differences caused by race of observer cannot be regarded as proven, because the Phase II study was not designed to permit unequivocal measurement of observer effects. However, as indicated in Chapter 5, Phase III data collection procedures have been altered to control race-of-observer effects and to provide better estimates of the magnitude of observer effects in general.

The next two sections discuss, in turn, the observation instrument used to describe the behavior of caregivers and the instrument used to describe the behavior of children. Instruments used to evaluate outcomes for children and parents are discussed in the final two sections.

OBSERVATIONS OF CAREGIVERS AND THE CLASSROOM ENVIRONMENT

In recent years educational researchers have turned increasingly to so-called "process-product" studies. That is, instead of merely evaluating the effects of the ingredients of educational programs — funds, curricula or equipment — on children's achievement, they have focused on actual classroom events (processes) that mediate between ingredients and achievements (products). The NDCS has adopted this approach in its evaluation of effects of center characteristics and, for this reason, extensive observations of caregivers were included in the study's design from the outset.

Phase II adult observations took place between January 15 and March 15, 1976. One hundred fifty-seven teachers and 90 aides were observed in 137 target classrooms. Most caregivers were observed for at least a full morning — approximately three and one-half hours. No caregiver was observed less than one-half hour. In classrooms without aides, the lead caregiver was observed on two mornings. In classrooms with aides, the lead caregiver and one or two aides were observed for one morning each.

The system selected for observing caregivers and the classroom environment, termed the *Adult-Focus Observation Instrument,* had been used previously by SRI in evaluating the Follow Through and Head Start Planned Variation Projects. The instrument's coding system was modified for the NDCS to record adult behavior in day care centers. The instrument has three parts — the Physical Environment Inventory, which describes the space, equipment and materials in the classroom; the Classroom Snapshot, which records numbers of staff and children present in the classroom and the general configuration of activities immediately preceding each Five-Minute Interaction observation; and the Five-Minute Interaction, a continuous record of the behavior of a particular caregiver.

Physical Environment Inventory

The Physical Environment Inventory consists of checklists noting the presence or absence of different kinds of equipment and materials, and ratings of various qualities of the environment (such as adequacy of space and equipment).

Types of equipment most commonly observed in target classrooms were dramatic play props, games, puzzles, reading materials, blocks and Legos, art materials and child-sized furniture. Those least commonly observed were live animals, work tools and skill equipment. Seattle classrooms tended to be better equipped than those in Atlanta and Detroit. Classrooms serving children from relatively high-income families tended to be better equipped than those serving low-income children.

Space or equipment problems were occasionally observed in half of the target classrooms. However, only 20 classrooms (15 percent) were consistently judged to have space or equipment problems. Paradoxically, Seattle classrooms were more often observed to have such problems than classrooms in Detroit and Atlanta.

Following the work of Prescott,[3] who found that *softness* or malleability and flexibility of objects in the physical environment is associated with positive outcomes for children, a *softness* index was computed, based on the presence of materials such as rugs, grass and water. *Softness* indices were higher in Seattle than in Detroit and Atlanta, and were higher in centers with high average family incomes and lower in centers with low average family incomes.

Classroom Snapshot

The Classroom Snapshot includes a count of adults and children, and a general description of their activities. Activities are coded into 24 categories. Each snapshot allows the observer to record multiple activities simultaneously. Table 3-2 shows the distribution of the activity codes recorded in the Classroom Snapshot. The table also shows how distribution of teachers and aides differed in the activities. Percentages shown in the table are based on a total of 3700 Snapshots of the 137 target classrooms.

*Receptive activities** for children – adult-directed activities, including stories, music and instruction in cognitive and practical skills – were recorded in 42 percent of all snapshots. *Exploration and use of materials* occurred in 45 percent of the snapshots. While a large number of codes were frequent, no one code was recorded in more than 20 percent of all snapshots. Among activities occurring least often were unusual creative activities and instruction in cooking and crafts.

For most of the activities observed, teachers were involved with children more often than were aides. Aides were most frequently observed handling arrival and departure of children and dealing with uninvolved, wandering children.

The snapshots also provided information on relatively infrequent events that may be of critical significance because they have an important positive impact on development or because they signal potential harm to children. Crying children and children being disciplined or comforted by caregivers were all relatively rare. While crying was observed in fewer than five percent of observations in the average classroom, in a few classrooms crying occurred in as many as 70 percent of the observations. Disciplinary incidents were observed in seven percent of snapshots overall but in a few classrooms appeared in almost every snapshot. Comforting of distressed children occurred in three percent of the snapshots.

Five-Minute Interactions

The Five-Minute Interactions are detailed records of the interactions of a selected caregiver. They take the form of "sentences" identifying an actor *(who)*, the object of the action *(to whom)*, the nature of the action *(what)*, and the style of the action *(how)*. During each five-minute observation period,

---

***Receptive activities** *and* **Exploration and use of materials** *refer to a convenient grouping of codes. They do not label a variable being used analytically.*

Table 3-2
Relative Frequencies of Classroom Snapshot Activities
(Based on 3700 Snapshots of 137 classrooms)

Activity	Average Percent of Snapshots in which Code was Recorded[a]	Percent of Occurrences with Teachers Involved in Activity	Percent of Occurrences with Aides Involved in Activity
Child being disciplined	7	61	39
Child arriving/departing	5	43	57
Child being nutured/comforted	3	67	33
Uninvolved child	11	44	56
Child observing	5	45	54
Receptive activities [b]	*42*	*72*	*28*
Story reading	10	68	32
Teacher-directed music, dancing	9	71	29
Cognitive skill instruction	10	68	32
Cooking, carpentry instruction	1	75	25
Practical skill instruction	12	65	35
Expressive activities [b]	*13*	*74*	*26*
Group sharing	5	80	20
Social interaction	8	68	32
Exploration and use of materials [b]	*45*	*64*	*36*
Table games	11	64	36
Lego, peg board (limited exploration)	14	50	50
Unusual creative activities	2	62	38
Sand, dirt, collage (open exploration)	18	66	34
Dramatic play	14	59	41
Active play	15	62	37
Doing work/helping [b]	*41*	*66*	*34*
Setting table, clean-up	5	66	34
Lunch	19	69	31
Transition	17	63	37

---

[a] Figures in this column represent the average percent of snapshots in which each activity appeared; percentages were first averaged across snapshots for each classroom and then averaged across classrooms.

[b] The frequency of each of the indicated categories is the sum of its component activities.

up to 63 sentences can be recorded describing the behavior exhibited by and directed at the focal caregiver. Examples of codes in each category appear in Table 3-3, which shows how often caregivers were observed in various types of interaction (such as questioning, correcting) and how often caregivers interacted with other adults, individual children and groups of various sizes. The table also shows the frequency of the more common stylistic modifiers (such as touching or dramatic play) when caregivers were dealing with one or two children. The total number of Five-Minute Interactions recorded was 3715, or an average of 15 for each of the 247 observed caregivers.

The caregiver behaviors most frequently observed were *center-related activity* (25 percent of all observations) and *observes/attends* (26 percent). *Center-related activity* was a general category used to describe verbal and nonverbal interchanges involving center equipment or procedures. Examples include a

Table 3-3

Examples of Relative Frequencies of Caregiver Behavior Codes

"To Whom" Codes		"What" Codes		"How" Codes[b]	
Code	Average Percent[a]	Code	Average Percent[a]	Code	Average Percent[a]
Adult	8.3	Command	11.9	Touch	7.2
One/two children	57.8	Question	10.3	Nonverbal	3.9
3 - 7 children	8.7	Respond	1.7	Positive/Happy	2.7
8 - 12 children	9.5	Instruct	6.2	Negative/ Threatening/ Punitive	0.9
13 or more children	15.0	Adult-related activity	2.2		
		Center-related activity	25.2	Guide to Alternative	0.9
		Comfort/support	1.2	Rule	0.5
		Praise/acknowledge	5.4	Dramatic play	0.2
		Correct	12.6	Materials	5.0
		No response/reject	0.3		
		Observe/attend	25.8		

[a] Each percentage represents the proportion of all observations in which the actor ("who") was a caregiver (teacher or aide).

[b] "How" codes are shown only for situations in which the caregiver was interacting with one or two children. "How" codes were optional and were not recorded for all observations.

teacher cleaning up, commenting that "swings (present in the center) are fun," or directing children to take out coloring books. *Observes/attends* refers to monitoring of children's activity and was usually recorded when the caregiver was in the presence of groups of children.

The most common caregiver behaviors with individual children were *corrects, commands* and *questions. Instruction* occurred in a somewhat smaller proportion of observations. *Correcting* was frequently accompanied by "how" codes indicating negative affect *(negative/threatening/punishing).* Other forms of interaction were generally accompanied by positive affect, if any affect was recorded at all. The relative infrequencies of the affect codes may be explained by their strict definitions, which were necessary for purposes of inter-rater agreement. No affect was recorded unless clear behavioral indications of feeling were present. For example, positive affect was recorded only when the caregiver used courteous language, accompanied by smiling. Physical contact occurred in seven percent of interactions between caregivers and individual children or pairs of children.

In addition to the behavior codes recorded directly on the Five-Minute Interaction, two constructs based on collections of codes were investigated; these were labeled *warmth* and *interactiveness.*

Warmth is one general dimension of teacher behavior that has consistently shown positive effects in research on elementary education.[4] A dimension of *warmth* was constructed from the adult observation data by combining three codes – the "what" codes labeled *praise* and *comfort/support,* and the "how" code labeled *positive* (referring to overt indicators of positive affect, such as politeness and smiling). The codes comprising *interactiveness* were grouped on a common sense basis and indicate the caregiver's attempt to stimulate positive actions on the part of children. The construct includes the codes labeled *praises, questions, instructs* and *corrects. Warmth* was coded in 12 percent and *interactiveness* in 35 percent of all observations.

Teachers and aides showed different patterns of behavior, reflecting their different roles in the classroom. Teachers more frequently asked direct questions and instructed children, while aides were more likely to observe. Teachers and aides offered praise and support with equal frequency, and displayed equal amounts of positive and negative affect. Teachers had significantly higher scores on *interactiveness,* while there was no difference in *warmth* between the two groups.

Frequencies of most caregiver behaviors differed from site to site. Generally, caregivers in Seattle diverged from those in Detroit and Atlanta, showing greater frequencies of *correcting, praising, positive affect* and *warmth.* Detroit caregivers, however, showed greater frequencies of *support and comfort.*

OBSERVATIONS OF CHILDREN

Outcomes of day care and early education programs have usually been measured with standardized tests of cognitive, socioemotional and physical development. Relatively few studies have used direct, naturalistic observations of children, either to describe the social environment of the day care center or to assess the impact of centers on children's behavior. The observations of children conducted as part of the NDCS were designed to serve both purposes.

During Phase II, 1244 children were observed in October-November 1975, and 1350 in May 1976. In the fall, each child was to be observed for three 20-minute periods, once at mealtime, once in a teacher-directed activity and once in free play. Because of scheduling difficulties the mealtime observation was dropped and replaced with a second free-play observation in the spring.

The system used for observing children, the Prescott-SRI Child Focus Observation Instrument, was adapted from one previously used to study day care by Elizabeth Prescott and her colleagues.[5] The instrument was further modified on the basis of the fall experience. The modified Prescott system consists of 20-minute observation periods broken into 80 15-second coding intervals. Observers were provided with timers that clicked every 15 seconds and were instructed to record what was happening to a selected focus child at the time of each click.

The instrument describes the child's activities in terms of a set of fine-grained behavior codes. Table 3-4 shows the activity codes used in the revised instrument and their relative frequencies of occurrence in the spring observations. In addition to the activity codes, the *object* of the child's activity was coded (adult, child, a group, or the environment). Finally, an index of *activity continuity* indicated new activities, ongoing activities, return to previous activities, and noninvolvement in activities. A total of 4048 observations were made on the 1350 children in the spring sample. Since each observation consisted of about 80 separate coding intervals, a total of about 323,000 codes of each type was available for analysis.

In Phase II, the Child-Focus instrument was used to measure seven of the child variables listed earlier in this chapter — (1) *patterns of social interaction;* (2) *interest/participation;* (3) *apathy/withdrawal;* (4) *cooperation/compliance;* (5) *anger/hostility;* (6) *dependency;* and (7) *task persistence.* The first variable describes the general configuration of the child's social activities. The next four are variants of behavior clusters that have emerged repeatedly in studies of preschool children in group care settings.[6] These variables have been shown to be related to children's later social adjustment and cognitive achievements. Of the two remaining variables, *dependency* has a long research history in developmental psychology.[7] Finding mature, socially acceptable ways to handle dependent needs is one of the preschool child's most important tasks. *Task persistence* has a less extensive research history; it is theoretically related to the child's ability to function in school and shows a dramatic increase between the preschool and elementary school years.[8]

Relative frequencies of all constructs are shown in Table 3-4; constructs are defined in the following paragraphs. Codes included in each construct are indicated in the table. In reading the variable definitions and descriptive results that follow, it should be kept in mind that the variables reflect not only traits of individual children but the opportunities and constraints of the classroom situation as well. For example, some children are more hostile than others, but some situations (generally, frustrating ones) tend to elicit hostile behavior in most children. Thus a reduction in the level of hostility over the day care year can mean either individual children have become less hostile, the classroom has become less frustrating, or both.

Patterns of Interaction was a set of variables representing the frequency of children's interactions with adults, individual peers, groups of peers and the environment. Based directly on the object codes of the Prescott instrument, these variables described the distribution of the child's activities at a relatively gross level, providing a context for interpreting the frequencies of other variables. These variables were also used to determine whether high staff/child ratios and small group sizes in fact led to more frequent interaction between adults and children, as the FIDCR appear to assume they would. In this respect, the variables served to cross-validate similar information from the Adult-Focus Instrument. Finally, interaction with peers was of interest because of its role in developmental theory as a precondition for cognitive and moral growth.[9] Children in the NDCS sample showed a steady increase with age in *social interaction* (computed as the sum of all interactions with people as opposed to the environment); this finding is congruent with previous research.[10] Girls

Table 3-4
Frequencies of Child Observation Codes and Variables
(Spring 1976)

A. Activity Codes and Associated Variables

	Average Percent[b]
Interest/Participation[a]	**43.0**
Open ended, expressive activity	5.8
Considers, contemplates	0.5
Adds prop, idea	0.3
Acts creatively, solves problem	0.1
Offers to help, share	0.5
Shares open ended activity	17.2
Defends rights	0.3
Moves with purpose	4.7
Selects activity (alone)	0.5
Selects activity (with others)	0.7
Asks for information	0.7
Asks permission to share	0.3
Gives opinions	8.2
Asks for recognition	0.9
Directs others in task	1.2
Intrudes playfully	1.1
Apathy/Withdrawal[a]	**8.8**
Wanders aimlessly	1.8
Minimal attention	1.0
Mimics, follows along	4.8
Ignores, no response	0.7
Avoids, withdraws	0.1
Cries	0.2
Isolates self	0.2
Dependency[a]	**0.5**
Asks for assistance, help	0.1
Asks for comfort	0.1
Receives comfort	0.3

	Average Percent[b]	Percent of Appropriate Responses[c]
Cooperation/Compliance[c]		**80.7**
Receives demands, requests	3.8	81.0
Receives invitation to play, share	0.6	65.2
Receives information	1.8	88.5
Receives guidance to task solution	0.3	91.8
Receives praise	0.2	81.6
Receives general comments, questions	4.9	86.2
Receives rules, corrections	0.4	72.2
Receives punishment, threats	0.2	48.5
Anger/Hostility[a]	**0.5**	
Shares hostile experience	0.1	
Intrudes hostilely, bullies	0.2	
Refuses to comply	0.2	
Hostilely asserts rights, anger	0.0	
Other Activity Codes Not Included In Constructs[d]		
Monitors environment (looks, watches)	10.3	
Closed, structured activity	2.5	
Shares closed structured activity	16.9	
Receives accidental intrusion	0.1	
Receives playful intrusion	0.9	
Receives hostility, being bullied	0.2	
States rules, tattles	0.3	
Offers sympathy, comfort	0.1	
Passive attention to activity	0.2	

B. Object Codes: Patterns of Interaction

Attention to adult	17.7
Attention to child	21.9
Attention to environment	50.1
Attention to group	9.8

C. Activity Continuity Codes: Task Persistence

Longest activity[e]	38.6

(continued)

Table 3-4 (continued)

---

[a] Each of the indicated variables is the sum of its component codes.

[b] Code frequencies are shown as a percentage of all observation codes. For both activity and object codes, the total number recorded was approximately 323,000.

[c] The "receives" codes indicate initiatives by others toward the child. The column headed "Average Percent" shows the frequencies of these codes as percentage of all 323,000 codes. The column headed "Percent of Appropriate Responses" indicates how often children responded appropriately to each of the various kinds of initiatives. *Cooperation/Compliance* was computed as the average percentage of appropriate responses.

[d] Seven codes, each of which occurred fewer than once for each thousand codes, have been omitted from this table. They include: *quits activity after frustration, angry reaction to frustration, temper tantrum, receives physical punishment, receives accident, receives/experiences pain, intrudes accidently.*

[e] The *task persistence* variable, or *longest activity* code, is computed as a percentage ratio of the duration of the longest activity to the total duration of the observation period. Since the observation period usually lasted 20 minutes, the longest activity of the typical child lasted .386 x 20 minutes, or 7.7 minutes.

showed a slightly greater amount of *social interaction* at all ages than did boys.

Interest/participation reflected the child's active involvement in the educational and social activities of the center. A large number of codes, shown in Table 3-4, were incorporated into this construct on the basis of theoretical considerations and examination of relationships among codes. There were no significant sex differences in levels of *interest/participation*. Older children showed higher levels than did younger ones. Children high in *interest/participation* also tended to score well on tests of cognitive achievement, described below.

Apathy/withdrawal reflected the child's tendency to separate herself or himself from the educational and social activities of the center. Predictably, *apathy/withdrawal* showed a strong negative relation to *interest/participation.* Older children exhibited less apathetic and withdrawn behavior than did younger children. Boys and girls did not differ in the amount of *apathy/withdrawal* they exhibited. In addition, children who scored well on tests of cognitive achievement showed little *apathy/withdrawal* when observed in centers.

Cooperation/compliance reflected the child's tendency to follow adult directions and to behave according to adult expectations. *Cooperation/compliance* was computed as a ratio. The numerator was the sum of all active, appropriate responses toward adult direction; the denominator was the sum of all responses recorded, including active, appropriate responses and passive or inappropriate responses. Girls tended to show more cooperation than boys; children of different ages were equally cooperative.

Anger/hostility included rebellious and aggressive behavior, not positive, self-assertive behavior. Older children tend to show less *anger/hostility* than younger children, and girls showed less at all ages than did boys. (This sex difference is consistent with a mass of earlier research.[11]) Overt anger and hostility were relatively rare in Phase II centers, accounting for less than one percent of all codes in both the fall and spring. However, as noted earlier, because a time-sampling procedure was used, the prevalence of a code depends on the duration of the relevant behavior, as well as its frequency. Because hostile outbursts in young children are often brief, codes reflecting such behavior tend to be infrequent. Nevertheless, they merit careful analysis as potential indicators of stress and harm in the day care enviornment. In fact, the frequency of hostile behavior dropped sharply between the fall and spring, possibly indicating the positive, socializing effect of day care reported by previous researchers.[12]

Dependency referred to the child's tendency to seek comfort and assistance from adults. (Other dependent behaviors, such as seeking praise and attention, were not included.) *Dependency* declined with age and was more prevalent in girls than boys. Both findings have been noted by previous investigators.[13] In addition, *dependency* appeared to decline between fall and spring in the day care setting.[14]

Task persistence referred to the child's tendency to devote sustained attention to a particular activity. *Task persistence* was calculated as the ratio of the duration of a child's longest activity in a particular observation sequence to the total length of the sequence (normally 20 minutes). *Task persistence* increased with age – again a finding consonant with earlier research.[15] Girls showed a slight tendency to persist at tasks longer than boys.

In sum, the observation data provided a foundation for building a number of constructs of recognized psychological importance. The NDCS findings regarding distribution of these variables by age and sex paralleled those of previous research. Preliminary examination of effects of the policy variables on these constructs was conducted during Phase II. A more definitive analysis is underway in Phase III, as is development of new child behavior constructs.

TESTS AND RATINGS OF SOCIOEMOTIONAL AND COGNITIVE DEVELOPMENT

In addition to information on child outcomes available through direct observation, data on socioemotional and cognitive development were gathered with a battery of standardized tests and rating scales. During Phase II, the battery included the Preschool Inventory, the McCarthy Scale of Verbal Memory, the Matching Familiar Figures Test, the Motor Inhibition Test, and the Pupil Observation Checklist. Teacher ratings were also investigated as a potential outcome measure. All tests were administered in fall 1975 and spring 1976, with the exception of the Motor Inhibiton Test, which was administered in the spring only. A total of 1260 children were tested by SRI in the fall and 1437 in the spring.

The Preschool Inventory (PSI)

The PSI is a test of certain cognitive skills and knowledge.[16] Developed by Bettye Caldwell for the Educational Testing Service, it has previously been used in the Head Start Longitudinal Study and Head Start Planned Variation Study. The test was not designed as a measure of general intelligence, but as a measure of school readiness.

The PSI proved highly reliable for the NDCS population.* The test was sensitive to age; children gained about one point a month on the 64 item inventory. (Table 3-5 shows average PSI scores by age of child.) Children in the NDCS sample from low-income families scored substantially higher than children from comparable backgrounds in the two Head Start samples mentioned above.

Internal analyses of PSI items revealed that verbal skills could not be separated from other abilities, as had originally been hoped. A shorter version of the test, consisting of 32

**Internal consistencies for various subsamples hovered around .9; the fall-to-spring test-retest correlation was .87.*

items, can be used to measure general ability with little sacrifice in psychometric quality. (Scores on the short version correlate .96 with scores on the longer version.)

Table 3-5
Average Scores on the Preschool Inventory, By Age
(Maximum = 64 points)

Child Age	Average Score
37 - 42 months	24.2
43 - 48 months	30.2
49 - 54 months	38.0
55 - 60 months	43.6

McCarthy Scale of Verbal Memory (MVM)

The McCarthy Scale of Verbal Memory, as its name suggests, taps both language and memory skills.[17] It consists of six items: the first four are lists of three to four words and the last two are sentences containing seven and nine criterion words, respectively. Children are asked to repeat each item after it is read. Their scores are the numbers of words or sentences they repeat correctly. The McCarthy Scale, like the PSI, demonstrated excellent internal consistency.* However, the test showed a pronounced ceiling effect; that is, children's scores did not continue to increase after they reached about 20 points (generally between four and four-and-a-half years of age). Table 3-6 shows average scores on the McCarthy scale for children of different ages.

Table 3-6
Average Scores on the McCarthy Scale of Verbal Memory, By Age
(Maximum = 30 points)

Child Age	Average Score
37 - 42 months	14.7
43 - 48 months	16.9
49 - 54 months	19.7
55 - 60 months	20.9

Matching Familiar Figures Test (MFFT)

The MFFT is a test of cognitive style, designed to measure a child's reflectiveness or impulsiveness.[18] Children are asked to match pictures of commonplace objects; their error rates and speeds of responding are measured. Children who respond quickly but incorrectly are classified as impulsive, while children who respond slowly but correctly are classified as reflective. Although the test has been used successfully in research with older children, it did not prove to be reliable or valid for the NDCS preschool population. The relationship between error rates and speed of responding was weak, and reliabilities for the error score ranged from .49 to .61 for various age groups. (Other researchers have also recently questioned use of the test with preschoolers.[19])

Motor Inhibition Test (MIT)

The MIT was intended as a measure of self-control, specifically of the child's ability to control (inhibit) his or her motor responses. In one segment of the MIT, the child is asked to walk a straight six-foot path drawn on the floor and is then asked to walk the path as slowly as he or she can. (The ability to move slowly in a precise path is an indicator of motor control.) In another segment, the child must draw a line as slowly as he or she can. Though the MIT has been used in previous research with preschool children,[20] it showed unacceptably low two-week test-retest reliability for the NDCS population.

Pupil Observation Checklist (POCL)

The POCL consists of nine five-point scales designed to assess the following bipolar dimensions of child behavior:

**Coefficients ranged from .87 to .91 for various age groups.*

1. resistive – cooperative
2. shy – sociable
3. outgoing – withdrawn
4. involved – indifferent
5. defensive – agreeable
6. active – passive
7. gives up – keeps trying
8. quiet – talkative
9. attentive – inattentive

The POCL was used in the NDCS as an indicator of the child's state during testing. SRI's testers rated each child on each of three days of testing.

Items on the POCL tended to cluster into two groups. Children's ratings on items 1, 4, 5, 7 and 9 tended to vary together suggesting an underlying dimension of *task orientation.* Similarly, items 2, 3, 6 and 8 varied together, suggesting an underlying dimension of *sociability.* This clustering, which occurred in both the fall and spring, duplicated a similar clustering in the POCL data from the National Home Start Evaluation.[21] Thus, the POCL appears to tap two important dimensions of behavior in the test situation rather consistently. Both *sociability* and *task orientation* increased with age.

Caregiver Ratings

During June 1976, SRI field-tested four rating scales designed to allow caregivers to assess children's development in several socioemotional domains. It was hoped that ratings by adults who knew children individually would provide a rich source of data that could be used as an independent check on information provided by observations and standardized tests. Several of the rating scales were explicitly intended to tap the same developmental variables as those constructed from the codes of the Child-Focus Instrument.

Four scales were tested:

- *Schaefer Day Care Behavior Inventory.*[22] A 30-item, five-point rating scale with three subscales: a) *introversion/extroversion,* b) *task orientation/distractability,* and c) *considerateness.* Items are rated "almost always" through "almost never."
- *Kohn-Rosman Problem Checklist.*[23] A 49-item, three-point scale with two subscales: a) *apathy/withdrawal* and b) *anger/defiance.* Items are rated "not at all typical," "somewhat typical," and "very typical," with the latter two categories indicating possible existence of problem behaviors.
- *Kohn-Rosman Social Competence Scale.*[24] A 73-item, five-point scale consisting of two subscales: a) *interest/participation – apathy/withdrawal* and b) *cooperation/compliance – anger/defiance.* Items are rated "very often or always" through "hardly ever or seldom."
- *Vineland Social Maturity Scale.*[25] An 11-item scale that assesses self-help skills associated with toileting, eating, dressing and avoiding hazards. Items are rated in terms of whether the child "usually or habitually," "occasionally but not usually" or "rarely or never" performs them with success.

Ten pairs of caregivers from each NDCS site were selected for the field test. These 60 caregivers were trained by SRI, and subsequently each pair rated 10 children familiar to both raters.

Results of the field test indicated that teacher ratings could not be used as outcome measures in the NDCS. Many of the correlations between ratings given by pairs of teachers to children well-known to them were low; that is, teachers often disagreed on which children should be rated relatively high or low. Even where correlations were high (relative ratings were in roughly the same order), there were often large absolute differences in ratings assigned to individual children, suggesting that caregivers tended to use the various scale designations differently. Finally, when agreement between raters was high, the reason often was that almost all children received the same score on a particular scale. For example, 80-90 percent of children were seen as being free of the psychological difficulties identified in the Kohn-Rosman problem checklist, while virtually all children were seen as possessing the skills identified in the Vineland Social Maturity Scale. In sum, there was a clear danger that differences in caregiver ratings across centers would reflect differences in caregiver perspective as much as genuine differences among children. There was also a danger that some of the scales would not differentiate children adequately.

Summary

The Preschool Inventory appears to be a reliable indicator of school-related cognitive skills and, as such, will be a useful outcome measure in the NDCS. The Pupil Observation Checklist captures some important aspects of children's test-taking behavior; data from the POCL may aid in interpreting other test scores and in cross-validating some observation data, notably *task persistence* scores from the Child-Focus Instrument. Other measures of socioemotional and cognitive development have proved to be of questionable value for use with the NDCS population. Therefore, the latter tests have been dropped from the Phase III battery, while several new tests have been added, as discussed in Chapter 5.

PARENT INTERVIEWS

Parent interviews were included in the NDCS for three reasons. First, they provided measures of parent satisfaction – one of the most important outcomes of day care. Second, they provided information on the degree and nature of parent involvement in center activities; parent involvement is mandated by the FIDCR. Finally, they provided information on the child's background—basic demographic information, as well as information on parental attitudes and practices regarding the raising of children, parental expectations about day care, and the family's previous experience with day care. Because these background factors can influence child behavior and test scores profoundly, they must be taken into account in assessing the effects of regulatable characteristics of the day care center.

In fall 1975, interviews were conducted with 1145 parents of target children. In all cases, only one parent was interviewed; in almost all instances this was the mother. All parents provided the basic family background data discussed in Chapter 1.

In addition, parents of children new to the centers were asked about: (1) the family's use of day care (previous day care for target child and other children in the family, reasons for needing day care, reasons for selecting center care rather than family day care or babysitting, reasons for selecting the particular center); (2) their expectations regarding the center's program; and (3) their view of the importance of supervision and discipline, caregiver training, experience, skills and personality characteristics, as well as parental involvement in center operations. Parents whose children had been in the Phase II centers four months or more were asked about: (1) the nature and amount of their involvement in center operations; (2) their satisfaction with the center's cost, convenience, facilities, staff, program and opportunities for parent participation; and (3) their perception of the center's impact on their child, contribution to their own knowledge of child-rearing and related matters, and assistance with family problems.

Family Use of Day Care

Data on family use of day care are shown in Table 3-7. Most parents had some previous experience with day care, including center care, family day care homes and, less frequently, care by relatives inside and outside the child's home. Most parents cited the need to work as their reason for using day care, though some needed day care in order to pursue their education. A majority cited the superiority of educational programs for children as their reason for selecting center care over other forms of care. Others identified the presence of well-trained caregivers as a reason for preferring center care. Almost half of the respondents cited convenience of location, transit arrangements, or center hours as the major reason for their choice of a particular center. For some, the recommendation of a friend or relative was important. Few cited cost as a factor.

Parental Expectations

Table 3-8 shows what aspects of day care were most important to parents of children new to the centers, and what characteristics these parents wanted to see in day care staff. Parents believed strongly in the importance of trained and experienced caregivers. Yet they also believed that personal characteristics were more important determinants of caregiver quality than experience or education per se. They identified patience and understanding as the most important elements of the caregiver's personal style. In describing what they expected centers to do for their children, parents emphasized preparation for grade school, good supervision and discipline. They did not place as much stress on the child's need for freedom or individual attention. In general, parents did not expect to

Table 3-7

A. Previous Day Care Arrangements for Target Child*

Type of Care	Percent
In own home by relative	11
In own home by non-relative	12
In other home by relative	15
In family day care home	36
In center or nursery school	83

* Responses exceed 100% because many parents have experience with more than one kind of day care arrangement.

C. Reasons for Preferring Center Care to Other Forms of Day Care*

Reasons	Percent
Cost	7
Convenience	9
Physical facilities	6
Center staff	17
Educational program and services	71

* Responses exceed 100% because respondents have more than one reason.

B. Reasons for Use of Full-Time Day Care*

Reasons	Percent
So that I can work or look for work	74
So that I can attend school	17
To help my child learn things s/he needs for school	29
To teach my child how to be more social (sharing/cooperative)	27

* Responses exceed 100% because some respondents gave more than one reason.

D. Reasons for Selection of a Particular Day Care Center*

Reasons	Pub. Funded Ctrs. (%)	Parent-Fee Ctrs. (%)
Cost	16	18
Convenience	44	49
Recommended by friend or relative	21	32
Center Staff	4	14
Program and Services	18	35

* Responses exceed 100% because some respondents gave more than one reason.

Table 3-8

Parental Expectations of Day Care

A. Relative Importance to Parents of Various Aspects of Day Care

	Very Important (%)	Important (%)	Not Important (%)
Discipline	78	21	1
Allowing child to choose activities	33	63	3
Grade school preparation	80	18	2
Trained caregivers	90	10	—
Constant supervision	75	23	2
Individual attention	53	43	3
Caregiver experience	84	14	2
Parent participation in hiring directors	46	35	19
Parent participation in hiring teachers	27	45	28

B. Preferred Teacher Characteristics

Characteristic	Percent*
Personal style	92
Training and experience	37
Teaching technique	21

* Responses exceed 100 percent because respondents gave more than one answer.

play an active role in center decision-making; for example, less than half felt it was important for parents to help hire staff.

Parent Involvement in Center Activities

Of parents whose children had attended Phase II centers for at least four months, two-thirds were not actively involved in the centers. They visited only to confer with center staff, to observe their children or to attend social events. (See Table 3-9.) Twenty-three percent took advantage of the educational opportunities offered by the center through workshops, training sessions and parent education courses. A very small number of parents were employed at the centers or had a major role in decisions concerning the center. Although many parents wanted more involvement, virtually none were interested in an increased role in decision-making; instead, they voiced a desire to work as aides or to participate more actively in educational activities.

Table 3-9

Parent Activities at the Center

Activity	Current %	Preferred %
No involvement	24	6
Family problems (conference with director or teacher, social service or health personnel)	8	4
Observer or social activities	36	41
Volunteer — no decision-making (volunteer aide, help with kitchen, structural help, carpentry, painting)	5	26
Volunteer — some decision-making (Help in choosing activities, materials, planning meals)	2	6
Parent education (attend educational meetings, sit in on teacher workshops, training sessions)	23	12
Paid classroom aide	1	4
Major decision-making role (help to hire day care staff, buy equipment, etc.)	1	1

Parent Satisfaction

Parent satisfaction is an indicator of the extent to which centers met the needs of the population they served. Data on parent satisfaction are shown in Table 3-10. In general, parents were satisfied with their centers, particularly with cost of care, center location and center hours. However, parents also identified several areas in which centers could be improved. Some saw a need to strengthen and intensify educational activities of the centers. Some saw a need for more trained and experienced caregivers and for additional space. Almost half expressed concern about the safety of streets around the center, and 40 percent were dissatisfied with the center's arrangements for sick children.

Perceived Impact of the Center on Child and Family

Ninety percent of parents felt they had learned something about child care from the center. Information about children's needs, suggestions for handling discipline problems, and ideas about books, games and toys were among the specific kinds of learning cited. Most felt that their children's social behavior had changed for the better as a result of the center experience. Improvements were noted in language and other cognitive skills, as well as specific problem areas such as shyness, thumbsucking and separation anxiety. However, half of those interviewed also mentioned occasional bad behavior (usually unacceptable language) that they felt the child had brought home from the center.

Table 3-10

A. Satisfaction with Non-Programmatic Aspects of Center Care

Aspect of Care	Very Satisfied (%)	Some Concerns (%)	Very Dissatisfied (%)
Cost	66	31	2
Center hours	84	16	–
Center location	79	20	1
Center safety	47	49	4
Arrangements for sick children	59	40	1

B. Parent Satisfaction with Aspects of the Center Program

Aspect of Care	Very Satisfied (%)	Some Concerns (%)	Very Dissatisfied (%)
Discipline	64	36	–
Supervision	89	–	11
School preparation	61	35	5
Individual attention	89	–	11
Number of teachers	47	49	4
Number of trained teachers	35	57	9
Number of experienced teachers	37	56	7
Toys and materials	79	21	–
Center space	34	57	9

A third of the parents felt the center had helped them with money problems – in general, simply because the existence of the center allowed the mother to work. One in four received direct assistance from the center in getting a job and others received help in getting additional schooling or training. More than 20 percent received aid in dealing with some type of medical problem: parents of children in publicly funded centers consistently reported receiving more services of these kinds than those whose children were in parent-fee centers. (See Table 3-11.)

Implications

In general, Phase II centers met parents' practical needs and provided care that parents deemed adequate or better. Parent involvement in center activities and decision-making was not extensive, and while many parents desired increased involvement, few wished to participate further in making decisions. Two themes ran through many parents' responses to questions about their expectations, goals and evaluations of their day care arrangements: center care is seen as an educational opportunity, and staff characteristics are seen as important elements of quality care. Parents' views on those matters give added importance to staff qualifications as a policy variable and to school-relevant social and cognitive skills of the type being measured in the NDCS as outcomes of day care.

Table 3-11

Supplementary Assistance to Parents, by Percentage of Public Funding

Type of Assistance	Type of Center 0% Publicly Funded	1-25% Publicly Funded	26-75% Publicly Funded	76-100% Publicly Funded
With money problems	30	24	50	48
With medical problems	16	14	31	30
With job problems	24	20	32	28
In getting a job	30	14	24	30
In getting additional school or training	23	27	43	42

Chapter 4: NDCS FINANCIAL ANALYSIS

Two of the major questions the NDCS was designed to answer involve measurement and analysis of the cost of center-based day care:

- How is the per-child cost of center-based day care affected by variations in staff/child ratio, staff qualifications, group size and other regulatable center characteristics?
- How does the cost-effectiveness of center-based day care change when adjustments are made in regulatable center characteristics?

Detailed financial data will be collected from the 64 centers participating in the study for each of the 18 months from October 1975 to March 1977. These data will be used during Phase III of the study to develop a quantitative model of the determinants of day care centers' income and expenditures. By itself, the model will provide answers to the first question listed above. In conjunction with the results obtained from the NDCS analysis of classroom process and child outcomes, the model will also provide answers to the second question listed above.

This chapter presents a status report on the NDCS financial analysis — a description of the work completed by the end of Phase II of the study. The first three months of collection of financial data (October to December 1975) were used to implement and refine data collection procedures. Since the data reported here cover only the initial three-month shake-down period and six additional months beyond (through June 1976), the empirical results reported in this chapter are descriptive in nature and subject to revision when the full 18 months of data have been analyzed. At the end of Phase II, the accounting and economic equations which will comprise the quantitative financial model had been specified, but insufficient data were available to perform hypothesis tests and other statistical analyses.

The first section of this chapter briefly describes the Research Cost Accounting System used to gather detailed financial information from study centers. The second section presents a preliminary descriptive profile of the financial characteristics of NDCS centers, including average monthly cost and income per child. The third section describes the structure of the financial model which will be subjected to statistical testing during Phase III and used to analyze the costs and cost-effectiveness of center-based day care. The last topic discussed is the relationship between the financial analysis and the other components of the NDCS research — The National Day Care Center Supply Study and the analysis of child outcomes and classroom interactions.

RESEARCH COST ACCOUNTING SYSTEM (RCAS)

The RCAS was designed to collect monthly financial data from each study center according to standardized accounting categories of income and expenditures. The core of the data collection system is a Statement of Current Income and Expense, used to record cash income and in-kind donations, as well as cash and imputed expenses. Various worksheets were provided to centers to aid in the preparation of the income and expense statement — to record the quantity and value of donated goods and services, and depreciation expenses associated with purchased assets. To minimize the reporting burden imposed on centers and to ensure reliability of recorded information, data collection forms were tailored to the financial recordkeeping systems of individual centers.

The RCAS conforms to generally accepted accounting practices. Income and expenditures are recorded by categories which permit consistent comparisons across centers and across months of the year. Reporting forms permit computerized checks by NDCS staff for computational errors in recording information and for consistency of data received from one month to the next. Data from a sample of study centers for calendar year

1976 will be subjected to an on-site audit by an independent accounting firm during Phase III.

In implementing the RCAS, a number of adjustments had to be made to deal with problems specific to centers which were part of a larger organization or part of a chain of centers. In 10 of these centers, separate recordkeeping systems had to be set up to distinguish income and expenditures attributable to the individual center from the commingled financial records of the parent organization. In 24 centers, donated space was provided by sponsoring agencies. A consistent plan for valuation of this space was devised, using floor plans for the centers and an imputed rental rate based on local market values, and on the location and characteristics of the physical facility used by the center. In 18 centers, a portion of salaries of indirect administrative staff (e.g., a bookkeeper at the parent organization) had to be allocated to the individual study centers.

Financial data collected by the RCAS were linked on a monthly basis to program information collected from each center as part of the general data system for NDCS centers. The program information required to conduct the financial analysis included data on enrollment, attendance, staffing patterns and services provided, as well as salaries and background information on all center staff. Without these data, it would be impossible to determine causal links between center characteristics and financial status, especially cost of day care per child.

SUMMARY OF FINANCIAL CHARACTERISTICS OF NDCS CENTERS

This section presents an overview of average monthly income and expenses of NDCS centers for the period October 1975 to June 1976. The data reflect resource use rather than cash outlay, meaning that the cash value of donated goods and services and an estimate of depreciation are included in income and expense fingures. The results presented here are tentative. No data are included for summer months, when income and expenses may be atypical. In addition, the results, based on only nine months of data, may be revised when 18 months of data have been analyzed in Phase III.

Data on expenditures and income by site are presented in Table 4-1. NDCS centers operated at about the break-even level with an average monthly income of $8,913 and expenses of $8,672. Based on an average monthly enrollment of 79 full-time equivalent (FTE) children per center, average cost per child in study centers was $110 per month. There was, however, a sizable variation in cost per child among study centers. Although about one-half of the centers had an average cost between $51 and $100, nearly 11 percent averaged in excess of $200 per month. Among the three study sites, Atlanta had the highest average cost per child ($113 per month) and Seattle the lowest ($105 per month). Average cost per child in Detroit centers was $107 per month. The only cost category that showed a significant difference across sites was personnel expenses. Atlanta centers spent an average of $86 per month per child compared to $70 in Seattle and $68 in Detroit.

Cost per child was by no means homogeneous within sites. While the average monthly cost for Atlanta centers was $113 per child, 42 percent of Atlanta's centers averaged $100 or less, while 15 percent had an average cost in excess of $200. Similar variation was observed in Seattle and Detroit. Not surprisingly, personnel expenses accounted for the bulk of center expenses – $79 per child or nearly three out of every four dollars spent per month. Nine percent of monthly expenses ($10 per child) was spent for occupancy, and the remaining 19 percent ($20 per child) was spent on supplies, equipment and other operating costs.

Average weekly salary for center directors was $187 or about $9,700 per year. For teachers, the average was $113 ($5,850 per

Table 4-1

Average Monthly Income and Expenses by Site
October 1975 - June 1976

	Atlanta			Detroit			Seattle			All Centers		
	Total $	Per Child[a] $	%	Total $	Per Child[b] $	%	Total $	Per Child[c] $	%	Total $	Per Child[d] $	%
Average Income	$9,394	$111	100	$7,821	$113	100	$8,959	$116	100	$8,913	$113	100
Parent Fees	2,082	25	22	5,958	86	76	3,812	50	42	3,442	44	39
Government Payments	5,264	62	56	1,102	16	14	2,416	31	28	3,547	45	40
Gifts and Contributions	1,049	12	11	163	2	2	1,433	19	16	941	12	10
Other Income	999	12	11	598	9	8	1,298	17	14	983	13	11
Average Expenses	9,584	113	100	7,374	107	100	8,075	105	100	8,672	110	100
Personnel	7,340	86	77	4,670	68	63	5,421	70	67	6,214	79	72
Supplies	1,131	13	12	754	11	10	918	12	11	987	12	11
Occupancy	519	6	5	1,085	16	15	1,143	15	14	814	10	9
Other Expenses	594	7	6	865	12	12	593	8	8	657	8	8

[a] Based on average FTE enrollment of 85 children.
[b] Based on average FTE enrollment of 69 children.
[c] Based on average FTE enrollment of 77 children.
[d] Based on average FTE enrollment of 79 children.

year) and for aides $103 ($5,350 per year). Site differences existed at all three job levels: overall wages tended to be lower in Detroit, and teacher and aide salaries were higher in Seattle. Director* salaries were higher in Atlanta by about $25 per week, due to some extent to the relatively higher wage scale for directors in the Atlanta Public School centers.

Study centers received an average of $113 per child per month during the nine-month period. Government payments and reimbursement accounted for $51 (45%) of this income, with parent fees bringing in an additional $44. Gifts, contributions and miscellaneous sources accounted for the remaining $25 of monthly center income.

Study centers were almost evenly divided on the basis of primary sources of income (see Figure 4-1). Slightly more than half the centers were dependent on parent fees as their primary source of funds, the rest primarily on government payments. Average cost per child in parent-fee centers was substantially lower than in centers supported primarily by government payments ($88 vs. $140 per month). There was also much less variation in cost per child among parent-fee centers -- 75 percent of them had an average cost per child between $51 and $100 per month, and in only one center did this cost exceed $150. In contrast, only 22 percent of the centers dependent on government payments had an average monthly cost per child of $100 or less, while 25 percent had an average cost in excess of $200.

In Atlanta centers, public funds accounted for 56 percent and parent fees 22 percent of average monthly income. Parent fees, on the other hand, were the primary source of income in Detroit, accounting for 76 percent of average monthly income, while public funds represented only 14 percent. In Seattle, 42 percent of center income came from parent fees and 28 percent from public funds. In each case, miscellaneous sources such as gifts and contributions accounted for the remaining income.

Relative dependence on parent fees may in part explain the variation in cost per child. To a certain extent, centers receiving a relatively large share of their total income from government payments may be able to pass costs through to the government and may therefore be under less pressure to control costs. In contrast, market pressure may act as a restraint on cost in centers relying primarily on parent fees. This hypothesis will be rigorously tested when the NDCS cost model has been estimated.

THE NDCS FINANCIAL MODEL

During Phase II, work proceeded on the development of a model to estimate the financial effects of changes in regulatable day care center characteristics. The model consists of a set of interconnected equations, each one showing the relationship between a variable whose behavior is predicted by the model (dependent variable) and a set of explanatory variables. Among the key explanatory variables are the policy variables – staff/child ratio, group size and staff qualifications. Among the key dependent variables are average cost per child, rates charged to parents and the welfare department, and salaries paid to staff.

A preliminary list of the equations to be included in the model is presented in Table 4-2. Two general types of equations will be used:

- *Behavioral Equations:* These equations represent hypotheses about causal relationships among financial and program variables. These hypotheses will be subjected to statistical testing during Phase III and modified as the results of the statistical tests indicate. Once final versions of these equations have been developed, they can

---

Persons who function as directors in the Atlanta Public School centers actually have the job title "teacher." However, the NDCS has consistently treated these as directors in order to maintain compatibility with other centers which typically use job titles such as "director" or "supervisor."

Figure 1-1

Distribution of Centers by Primary Funding Source March Costs/March FTE Enrollment

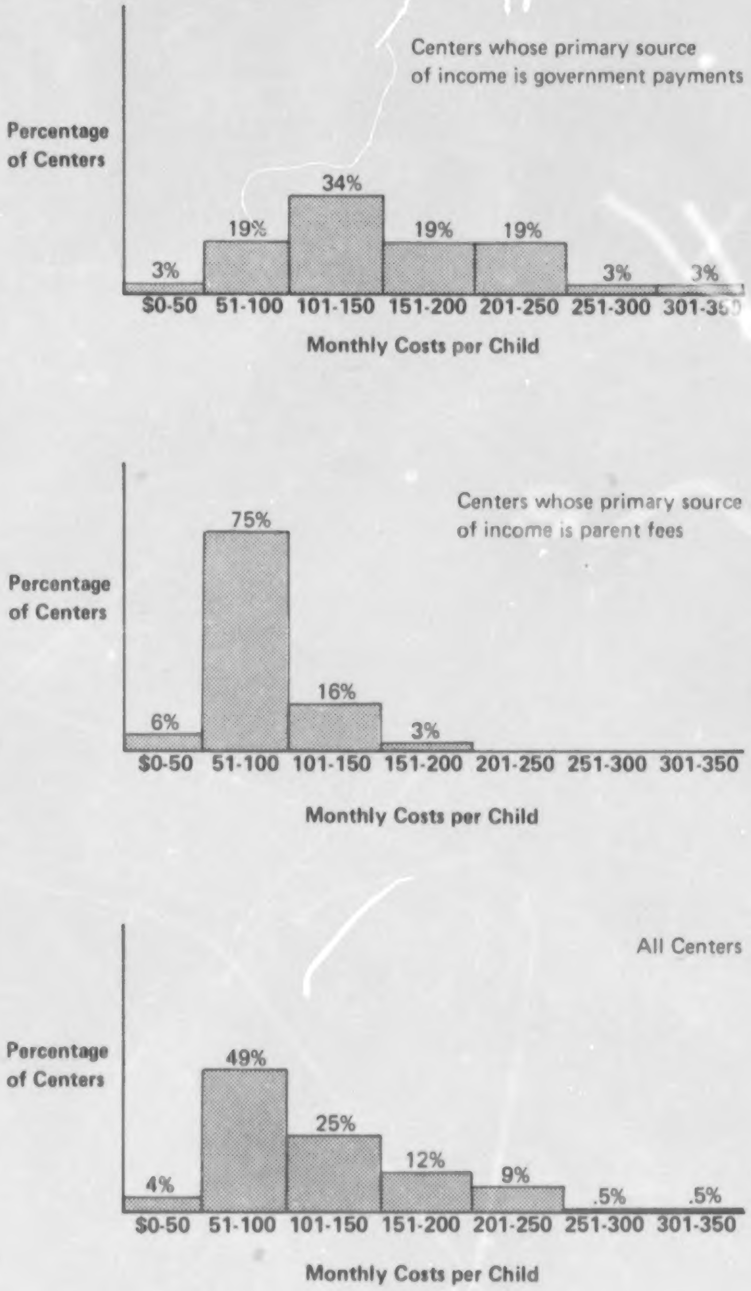


Table 4-2
Preliminary Specifications of the NDCS Cost Model

Behavioral Equations

1. DIRWAGE = f_1 (DIRED, DIREX GOVT, SITE)
2. GIFTS = f_2 (TKIDS, STATUS)
3. GKIDS = f_3 (COST, STATUS, SITE)
4. GOVFEE = f_4 (COST, STATUS, POS, SITE)
5. MIX = f_5 (STATUS, SITE)
6. OCCUP = f_6 (RSPACE, SITE)
7. OTHER = f_7 (TKIDS)
8. OTHNIC = f_8 (TKIDS, STATUS)
9. PARFEE = f_9 (COST, STATUS, ABS, UTILIZ)
10. PKIDS = f_{10} (PARFEE, SITE)
11. SERVE = f_{11} (STATUS, SITE)
12. SUPLS = f_{12} (TKIDS, STATUS, SITE)
13. WAGE = f_{13} (EDAT, EXPER, MIX, GOVT, SITE)

Accounting Identities

14. COST = EXPEND/TKIDS
15. EXPEND = PAROLL+SUPLS+OCCUP+OTHEXP
16. INCOME = TUIT+WELF+GIFTS+OTHNIC
17. NETERN = INCOME−EXPEND
18. PAROLL = STAFF x WAGE
19. RSPACE = TKIDS x SREQ
20. STAFF = SCR x TKIDS
21. TKIDS = PKIDS+GKIDS
22. TUIT = PARFEE x PKIDS
23. UTILIZ = TKIDS/CAP
24. WELF = GOVFEE x GKIDS

Variable Definitions

ABS = Dummy variable indicating whether or not the center charges for days absent

CAP = Licensed capacity

COST = Average cost per child

DIRED = Director's education attainment

DIREX = Director's day care experience

DIRWAGE = Director's wage

EDAT = Education attainment of center's staff

EXPEND = Total expenditures

EXPER = Day care experience of center's staff

GIFTS = Gifts and donations

GKIDS = Publicly subsidized enrollment

GOVFEE = Average government fee

GOVT = Proportion of center enrollment that is publicly subsidized

INCOME = Total income

MIX = Teachers as a proportion of center staff

NETERN = Net earnings

OCCUP = Occupancy expenses

OTHEXP = Other expenses

OTHINC = Other income

PARFEE = Average parent fee

PAROLL = Personnel expenses

PKIDS = Private enrollment

POS= = Purchase of service dummy variable

RSPACE = Required center area (sq. ft.)

SCR = Staff child ratio

SERV = Other services provided to children and families

SITE = Dummy variable indicating center's site

SREQ = Required space per child

STAFF = Total number of staff

STATUS = Legal status (profit or nonprofit)

SUPLS = Expenditures for supplies

TKIDS = Total enrollment

TUIT = Tuition income

UTILIZ = Center's utilization rate

WAGE = Average wage payment for center's staff

WELF = Welfare income

be used to forecast changes in each dependent variable which would result from a given set of changes in the determinants of each variable. One hypothesis to be tested, for example, is that the wage rate paid to center staff is dependent on staff characteristics (education, previous experience), characteristics of the center and locally prevailing labor market conditions (Table 4-2, equation 13). Because the precise forms of the behavioral equations are not known, these equations are shown in the table as unspecified functions, denoted by f_1, f_2 and so on.

- *Accounting Identities:* These equations reflect relationships among financial or program variables which are mathematical identities. For example, for all centers and for all monthly reporting periods, total center expenditures will always be equal to the sum of expenditure components (Table 4-2, equation 15). In addition, a center's capacity utilization rate by definition is equal to the number of children enrolled divided by the center's licensed capacity (Table 4-2, equation 23).

The remainder of this section discusses the most important of the equations to be included in the model and provides a justification for the causal hypotheses that the equations represent.

The first major dependent variable in the model is the wage rate for center employees. The director's wage rate is specified as a function of the director's educational attainment and day care experience, the proportion of the center's enrollment that is publicly funded, and the site in which the center is located. The average wage rate for other center staff is specified as a function of the average educational attainment and day care experience of the center's staff, the proportion of center staff whose primary job title is "teacher," the proportion of enrollment that is publicly subsidized, and the site variable. These specifications are consistent with those used in previous studies of the determinants of interoccupational wage differentials in the general economy. It is common practice in education and other industries for employers to base salary on level of education and previous experience. One would therefore expect a center's average wage payment to be related to the educational attainment and previous day care experience of its staff. Teachers are expected to receive higher salaries than aides. Centers whose staffs are more heavily weighted towards teachers should therefore have higher average salaries than other centers. The site variable is intended as a proxy for the effects of local economic conditions such as the prevailing wage structure in the local labor market.

Total expenditures by the center are broken down into personnel and nonpersonnel expenditures. Personnel expenditures are simply the product of the number of staff employed and the average wage paid by the center, including fringe benefits. Nonpersonnel expenditures are broken down into expenditures for supplies, occupancy and other items. Expenses for supplies should depend on the center's total enrollment, local prices and the center's legal status (profit or nonprofit). As before, the effect of local prices will be captured by a site variable. Occupancy expenses should depend on total enrollment, mandated space requirements and local unit cost (as measured by the site proxy variable). Other expenses are specified as a function of total enrollment.

The average fee charged to parents is specified as a function of cost per child, the center's utilization rate, the center's legal status, and a variable indicating whether the center charges for days absent. The primary determinant of the center's average fee is expected to be its cost per child. However, centers that charge for days absent are expected to have lower nominal fees than centers that do not charge for days absent. Centers that receive a significant proportion of their income from sources other than fees are expected to have a lower fee structure than centers that rely almost entirely on fees. Preliminary analysis indicates that nonprofit

76

centers receive approximately one-third of their income from gifts and contributions. The legal status variable is intended to pick up this effect. To the extent that a center finds demand for its services to be responsive to changes in its fee structure, its average fee may be dependent on its utilization rate. For example, a center that finds total enrollment falling short of licensed capacity may reduce its fees in an attempt to increase enrollment. Alternately, if a center believes that reducing fees will not attract additional families, it might respond to underenrollment by increasing fees in order to cover costs. In either case, fees would depend in part on utilization rates.

As in the case of parent fees, the cost per child is expected to be the primary determinant of the average fee paid by the government for publicly subsidized children. If the welfare agency guarantees payment for a certain number of slots, then the center operates under relatively little uncertainty and will be able to charge lower fees. In addition, the government may be in a position to obtain more favorable rates than the individual family. One would therefore expect centers that have purchase of service contracts with a welfare agency to have lower average fees than other centers.* The site variable is intended as a proxy for the effects of different rate-setting formulae and their institutional differences.

Total enrollment is the sum of private and publicly subsidized enrollment. Private enrollment is expected to depend on the fee charged to parents, while publicly subsidized enrollment is expected to depend on the cost per child and the center's legal status. The site dummy is included in both equations to reflect the demographic composition of sites.

Total income is the sum of its component parts – tuition, welfare fees, gifts and other income. Tuition is the product of private enrollment and the average fee charged to parents. Similarly, income from welfare fees is the product of publicly subsidized enrollment and the average fee paid by the government. Gifts and other income are each specified as a function of the center's total enrollment and its legal status.

Finally, the center's staff mix (teachers as a proportion of total staff) and the number of services provided to children and families are each specified as a function of the center's legal status and site. Profit-making centers are expected to employ relatively more aides and provide fewer services than nonprofit centers.

The income and expenditure variables in the NDCS cost model will be measured on an annual basis. Enrollment will be measured by the average full-time equivalent enrollment over the course of the year. The model will be estimated during Phase III, when 18 months of data will be available from the RCAS. The overall structure of the model will be re-examined during the estimation process. In the preliminary version, an attempt was made to avoid unnecessary complexity. As the estimation process progresses, relationships outlined in the initial version will be expanded and refined.

RELATIONSHIP OF THE FINANCIAL ANALYSIS TO OTHER COMPONENTS OF THE NDCS

The financial model described above will be linked to other components of the NDCS. The RCAS model is designed to predict changes in the financial status of day care centers (e.g., cost per child) which would result from a change in regulatable center characteristics. Data from the Supply Study

---

Earlier it was noted that federally funded centers have higher costs per child than parent fee centers. This differential may be due to the fact that federally funded centers maintain higher staff/child ratios than parent fee centers, or to differences in education or experience of staff in the two types of centers. The hypothesis suggested here is that, when differences in staffing patterns and other characteristics are held constant, centers with purchase-of-service contracts may have lower costs per child than other centers.

will be used to augment the RCAS model with additional behavioral equations. These equations will be used to forecast the changes in center characteristics which are likely to occur as a result of changes in government day care regulations*, monitoring procedures and reimbursement policies. The addition of these equations will complete the causal link between policy and administrative decisions, on the one hand, and changes in the characteristics and cost of center care on the other hand. Combined with NDCS findings on effects of variations in center characteristics on the safety and development of children, the complete financial model will provide the basis for the cost-effectiveness analysis to be included in the NDCS final report.

During Phase III, data will be available from the Supply Study on the program characteristics and financial status of over 3,000 day care centers, selected from the 50 states and the District of Columbia. Among these 51 geographical areas, there is a great diversity in local day care regulations, government reimbursement rates, day care monitoring practices and socioeconomic conditions. This diversity should be reflected by differences in the staffing patterns, program characteristics and financial status of the 3,000 centers included in the Supply Study sample. An important component of the statistical analysis of Supply Study data will be the development of behavioral equations which can be used to predict changes in center characteristics which would result from changes in government day care policies and practices — changes similar to the variations which exist presently among the 50 states.

By providing insight into relationships between government day care policy and center characteristics, the Supply Study adds an important dimension to the NDCS. For example, the adoption of a new version of the FIDCR with more lenient staff/child ratio requirements is likely to affect center characteristics other than staff/child ratio — especially in the absence of simultaneous changes in government child care reimbursement rates. Actual staffing ratios may not decline to the full degree anticipated in the new regulations. To the extent that personnel costs do decline, resulting changes in centers' fee structures may increase the utilization of FIDCR-certified centers by non-subsidized families and thereby increase the degree of socioeconomic integration in FIDCR centers. By examining both the direct and indirect consequences of changes in government day care policy, the NDCS will provide a thorough assessment of the relative merits of alternative policy decisions.

The RCAS and Supply Study models are complementary. The RCAS model, in conjunction with the results of the child effects study, is most useful in identifying which changes in day care centers are most desirable. The Supply Study model is most useful in identifying the combination of government policy decisions (specific regulations, reimbursement rates and monitoring practices) most likely to produce the desired changes in center programs. Together the two models provide an excellent means for translating policy objectives — the safety and development of children, reasonable cost economy, inter-person and inter-area equity, and program efficiency — into specific policy decisions and administrative action.

Note that a change in regulations is different from a change in regulated characteristics. Actual characteristics of centers may not fully reflect federal requirements. The degree of compliance will depend on monitoring procedures, penalities for noncompliance and so on.

Chapter 5: PHASE III DESIGN AND DATA COLLECTION

Phase III, running from October 1976 through January 1978, will provide the data on which most of the policy conclusions of the NDCS will ultimately rest. Phase II furnished valuable information on naturally occurring patterns of association among characteristics of centers (notably staff/child ratio, group size, staff education and staff experience), costs of care, and behavior of children and caregivers; it also provided useful knowledge about techniques for measuring these variables and analyzing the resulting data. In Phase III, the substantive and methodological insights of Phase II will be brought to bear on the policy questions identified in the **Introduction** to this report.

Phase III is a quasi-experimental[1] investigation in which selected aspects of centers – those which can be regulated – have been altered systematically, permitting measurement of the costs and effects associated with such changes. The Phase III design has two components. One is a study conducted in 49 centers in Atlanta, Detroit and Seattle in which staff/child ratios have been raised in a selected group of centers and left at their natural lower levels in a second group of centers. The two groups are matched with respect to a variety of staff and child characteristics. Included is a third comparison group of centers with naturally high ratios.

The second component of the Phase III design is a study conducted in 28 classrooms in eight centers operated by the Atlanta Public School (APS) system. In the eight-center experiment, children have been assigned randomly within centers to classrooms that vary systematically in staff/child ratio and staff education.

This dual approach has been adopted because the two designs have different strengths. The eight-center experiment has a potentially high degree of internal validity. Since the child population of the APS centers is relatively homogeneous in racial and socioeconomic characteristics and since children have been assigned randomly to different classrooms, the entering characteristics of children are unlikely to confound with ratio and staff education. On the other hand, the 49-center investigation, performed in three markedly different sites with a heterogeneous sample of children, is likely to yield results that are widely generalizable.

It should be stressed that the two components of Phase III are complementary, not separate, studies of the same policy issues. They will draw on identical sets of data concerning children's performance on standardized tests; child and caregiver behavior in centers; parent-center relations and parent satisfaction; center goals, programs, organization and operations; and costs. Clearly, convergent results from the two studies will constitute a far sounder basis for policy conclusions than would results of either in isolation. Divergent results will pose problems for both statistical and policy analysis, but this risk has been accepted in preference to the risk of presenting univocal but misleading information to the public and the policymaker.

Data collection plans are outlined in a final section of this chapter, following more detailed descriptions of the two designs.

THE 49-CENTER QUASI-EXPERIMENT

The purpose of this quasi-experiment is to provide broad-based, generalizable data on the impact of the major policy variables. Particular effort has been devoted to creating a design that will yield unambiguous causal conclusions regarding staff/child ratio by the addition of new caregivers in selected centers where ratios were low in Phase II.*

**The requirement that manipulation be accomplished only by making low ratios higher – that is, by adding caregivers – was specified by OCD in its original formulation of the design.*

Data on children, caregivers, parents, center operations, and costs gathered in these "treated" centers (Group I) are being compared to data from a closely matched comparison group of untreated, low-ratio centers (Group II) and from a second comparison group of untreated high-ratio centers (Group III). Entering characteristics of children, as measured by Preschool Inventory scores, have been carefully balanced across the treatment and comparison groups, as have group size and staff education. This design has been replicated in each of the three Phase II sites to permit assessment of site effects. To maximize the power of the design, all of the Phase II study centers, with only a few exceptions, have been retained in Phase III. Table 5-1 displays the basic features of the quasi-experimental design.[2] Key elements of the design and its implementation are discussed below.

Sites and Centers

Atlanta, Detroit and Seattle were retained as study sites in Phase III. Of the 64 Phase II centers, nine were public school centers in Atlanta and will be discussed in connection with the eight-center experiment described below. Six more centers were dropped, either because they closed, declined to participate, or proved to be extremely unstable or atypical during Phase II. The remaining 49 centers were retained for the quasi-experiment — 20 in Atlanta, 13 in Detroit and 16 in Seattle. No new centers were added because there

Table 5-1

Design for the 49-Center Quasi-Experiment

Group I = High SCR[t] (treatment)	HIGH PSI LOW PSI
Group II = Low SCR (low-ratio comparison)	HIGH PSI LOW PSI
Group III = High SCR (high-ratio comparison)	HIGH PSI LOW PSI

This design was replicated at each of the three sites: Atlanta, Detroit and Seattle. For each cell, treatment and comparison groups were balanced within site according to the projected Phase III center-average group size and the projected Phase III center-average years of staff education. All projections are based on Phase II data.

Key: PSI = Projected Phase III center-average PSI score
SCR = Projected Phase III staff/child ratio
SCR = Phase III staff/child ratio in treated centers. Ratios are fixed by contract. Staff are added to the designated centers to raise ratios from low values in Phase II to high values in Phase III.

seemed to be no obvious gaps in representation of center types in this existing pool.

Logic of the Design

As shown in Table 5-1, the treatment strategy partitioned the sample of centers into three groups:

Group I — Low-ratio centers to which staff has been added, making them high-ratio centers.

Group II — Low-ratio comparison group initially similar in ratio to Group I prior to treatment.

Group III — High-ratio comparison centers similar in ratios to Group I after treatment.

The treatment plan allows three comparisons to be drawn, each making a distinctive contribution to the analysis of the impact of staff/child ratio and other policy variables. The **Group I - Group II** comparison isolates the effects of ratio by a classical "experimental versus control group" design. The two groups are approximately identical, except for the fact that ratio has been altered experimentally in Group I. The **Group I - Group II** comparison is the only way to determine unambiguously whether changes in ratio can in themselves cause changes in the behavior of caregivers and children. Only by deliberately changing one group and leaving unchanged another otherwise similar group can the effects of ratio changes per se be isolated.

The Group I manipulation simulates one potential effect of the strengthened enforcement of the FIDCR under Title XX — namely an increase in ratios in centers serving publicly funded children but normally operating at ratios below those specified in the FIDCR. Note, however, that in the Phase III design the ratios were increased by an infusion of funds from outside the center; centers were not required to increase ratios by increasing their fees or siphoning off resources from other parts of their operations. Therefore, it is more precise to say the Phase III experiment simulates the effects of a public subsidy to centers for purposes of raising ratios to meet the FIDCR.

The **Group I - Group III** comparison (treatment centers/high-ratio centers) takes into account that high-ratio centers are likely to differ from those with low ratios in many characteristics other than ratio alone. Externally inducing an increase in ratio, whether through experiment or federal subsidy, may not cause centers to change in characteristics other than ratio. The **Group I - Group III** comparison will permit estimation of the effects of these characteristics on caregivers, children and families.

The **Group II - Group III** comparison (low-ratio/high-ratio) replicates Phase II, permitting an examination of the costs, center processes, and child outcomes associated with naturally high and low ratios, as well as with other characteristics of centers. Replication of Phase II associative findings will greatly enhance their credibility.

Use of Entering Child Test Scores as an Independent Variable

A major lesson learned in Phase II was that entering characteristics of children are intertwined with policy-relevant characteristics of centers in the day care world. For example, as indicated in Chapter 1, low-income children are concentrated in federally subsidized centers, which maintain high staff/child ratios. Low-income children also are served by relatively experienced caregivers.

Because SES and other characteristics of children are likely to show strong relationships with their behavior and test scores, assessment of the effects of the policy variables requires that SES be taken into account. During Phase II, this was accomplished by statistical procedures as described in Appendix A. During Phase III, an experimental approach is adopted specifically to separate the influences of child and center characteristics.

To make the comparisons as effective as possible, a key child status measure – the Preschool Inventory (PSI) score – was balanced across the three groups within each site. This use of child test scores as an independent variable is somewhat unusual and deserves explanation. Traditionally, measures of SES have been used to insure equivalence between comparison groups of children in educational and developmental research. Mother's education and family income, two common SES measures, could have been used as independent variables in Phase III, but SES does not fully guarantee equivalence of groups of children. At any level of family income or parent education, children of a given age will show a wide range of cognitive and social skills. A direct measure of those skills will produce a far more relevant form of group equivalence in a study using dependent measures that are likely to be sensitive directly to children's skills and only indirectly to SES. Outcome measures of the NDCS – child behavior and child test score changes (gains the child makes over time) -- are dependent measures of this kind.

It was therefore essential to control the best possible measure of the child's entering skills rather than SES which influences those skills but leaves a great deal of room for variation across families and individual children. As demonstrated in Chapter 3, the PSI is, on psychometric grounds, the NDCS's best single measure of the child's entering skills. Use of PSI does not indicate, however, that cognitive outcomes are viewed as more important than socioemotional ones.

Implementation of the Design

For Phase III, centers were assigned to cells in approximate accord with the design requirements shown in Table 5-1. Within sites, centers were assigned to treatment and comparison groups in such a way that staff/child ratio and PSI were independent of one another (and therefore will not be confounded in Phase III analyses). That is, within a particular ratio level at a given site, PSI exhibited almost the full range of variation in scores observed across all centers at that site during Phase II; similarly, within each level of PSI, ratio varied widely. Reasonable balance of staff education and group size was achieved across cells of the design, where cells are defined by site and by particular levels of ratio and PSI.

Average staff/child ratio in untreated centers prior to manipulation was projected to be approximately 1:9 (or .11). Average ratio after treatment was 1:6 (.17), an average increase of .06. The average change exceeds the standard deviation of naturally occurring ratios in Phase II (.05); moreover, the smallest change implemented in any center (.04) is very close to this value. In addition, estimated pre-treatment ratios fail to comply with the FIDCR, while ratios after treatment are approximately those mandated by the FIDCR.

THE EIGHT-CENTER EXPERIMENT

In addition to the 49-center quasi-experiment described above, Phase III also includes an eight-center experiment with the same general goals: providing evidence about the impact on children of changes in caregiver qualifications and staff/child ratio and providing data on the interactions of these two policy variables. Special attention has been given to the effects of the caregiver's level of education. The experiment is being conducted in eight centers operated by the Atlanta Public School (APS) system.

The design's independent variables, illustrated in Table 5-2, include staff education (three levels), staff/child ratio (two levels), and child age (two levels). Children of a given age (three or four years) are assigned randomly within each center to classrooms in which ratio and staff education vary systematically and independently, making it possible to assess main effects and interactions of staff education and staff/child ratio for children of different ages. Group size and caregiver experience have been balanced across cells of the design as far as was feasible within the limits of existing enrollments, physical facilities and staff resources. The dependent variables will be identical to those of the 49-center quasi-experiment.

The following subsections review the special characteristics of the APS centers that made the experimental design feasible and discuss the implementation of key features of the design.

The Atlanta Public School Centers

APS day care centers participating in the NDCS primarily serve publicly subsidized children, a principal policy focus of this study. These centers were remarkably similar in several respects, which recommended them for use in this design:

- Staff fell into three distinct categories, each marked by a particular educational

Table 5-2

The Eight-Center Design

Three-Year-Old Classrooms:

		Staff Education		
		High	Medium	Low
Staff/ Child Ratio	High (c.1:5)	N = 2 classes	N = 2 classes	N = 2 classes
	Low (c.1:10)	N = 2 classes	N = 2 classes	N = 2 classes

Four-Year-Old Classrooms:

		Staff Education		
		High	Medium	Low
Staff/ Child Ratio	High (c.1:5)	N = 2 classes	N = 4 classes	N = 4 classes
	Low (c.1:10)	N = 2 classes	N = 2 classes	N = 2 classes

"High" staff education:	the lead teacher is a center director, usually with a master's degree
"Medium" staff education:	the lead teacher is a graduate of Atlanta Area Tech two-year program
"Low" staff education:	the lead teacher has not received the Atlanta Area Tech degree

background. Center directors*, who are contracted to work in classrooms, all had bachelor's degrees, and most had master's degrees. Lead caregivers were mostly graduates of Atlanta Area Tech's two-year college training program for day care. Aides generally had high school diplomas or less. It was possible, by controlling assignment of caregivers, to create classrooms with clear contrasts in staff education.

- Children served by the APS centers were unusually homogeneous in ethnic and socioeconomic background. Therefore, effects of staff education and staff/child ratio would not be seriously confounded with effects due to demographic differences among groups of children in different classrooms.
- Many of the APS centers were large enough to contain two or more classes of children at age three, age four, or both. This circumstance made it possible to assign children randomly to different classrooms within the same center.
- Atlanta Public School authorities were extremely cooperative with the NDCS and agreed to reassignments of staff and children necessary to implement the design.

To meet design requirements, five centers were chosen from the Phase II group and three from other APS centers.

Implementation of the Design

Not only were the selected APS centers structured conveniently for our experimental design, but additional circumstances made implementation of the design relatively natural in the APS setting. Contrasting levels of the two major independent variables — caregiver education and staff/child ratio — were established as follows:

Caregiver Education: APS directors are contracted to work in classrooms, and aides are sometimes in charge of classrooms when fully qualified group leaders are unavailable. Use of both directors and aides as group leaders is within the scope of normal APS practice and made it possible for the NDCS to create classrooms in which caregivers differed appreciably in average education level without disrupting normal operations. Classrooms designated as "low" in staff education are being led by caregivers who have not completed the AAT course; classrooms with "medium" levels of staff education are being led by caregivers with the AAT degree; classrooms "high" in staff education are being led by center directors, most of whom possess master's degrees.

Staff/Child Ratio: During Phase II, the APS centers maintained high staff/child ratios, ranging from 1:4 to 1:8. Initially, it appeared that, given this range, a Phase III manipulation of staff/child ratio would not be feasible. To increase some ratios by adding another caregiver did not seem to be policy-relevant and to decrease certain ratios by removing one or two caregivers was not possible within the study's contractual constraints. APS personnel, however, felt that a ratio manipulation was, in fact, feasible because of the need for additional staff in other programs in the same schools. Transferring certain caregivers to these programs was not only possible but desirable from an APS point of view. Examination of child and staff enrollment data indicated that ratio contrasts of 1:5 and 1:10 could reasonably be obtained in the eight study schools; this contrast was incorporated into the experimental design.

Other Center Characteristics: Staff experience and group size were fairly well balanced across cells of the design shown in Table 5-2. Center averages in experience ranged from four to eight years across the 12 cells, with all but three cells falling between five and seven years. Though differences across cells were large in absolute terms, all of the centers fell in such a high range that their

---

**As noted earlier, persons who functioned as directors in the APS centers had the job title "teacher." Persons whose duties were comparable to those of "teachers" in most NDCS centers had the job title "group leader."*

84

variation seemed unlikely to affect processes and outcomes. Group sizes ranged from ten to 18 across the cells, with all but two falling in the range from 12 to 16.

PHASE III MEASUREMENT AND DATA COLLECTION

The Phase III data collection effort began in both the APS centers and the 49-center sample in October 1976 and is scheduled to end in spring 1977. As in Phase II, the Stanford Research Institute (SRI) is responsible for administering tests to children and observing both caregivers and children, while Abt Associates Inc. (AAI) is responsible for recording information on center staffing patterns, staff and child backgrounds, and center events and services, including staff and child absences, accidents, parent meetings, and staff training. All measurement and data collection instruments are identical for Phase III centers, with distinctions between the two components made only at the point of analysis.

The Phase III data collection plan consists of the following elements to be discussed further below:

- Test Battery (Fall 1976 and Spring 1977)
- Child-Focus Observation (Fall 1976 and Spring 1977)
- Adult-Focus Observation (Fall 1976 and Spring 1977)
- Center Organization, Operations and Cost Data (ongoing throughout the study year)
- Periodic Center Information (twice during year)
- Parent Interview and Survey (Spring 1977)
- Day Care Forces Inventory (Spring 1977)
- Site Case Studies (Spring 1977)

The first round of Phase III testing was conducted in October 1976. The second round will be completed in the spring of 1977 and during this testing, race of testers and children will be as closely matched as possible.

The following test battery is being used in Phase III:

Day 1 – The 32-item PSI and Pupil Observation Checklist (POCL)

Day 2 – The Peabody Picture Vocabulary Test (PPVT), a motor scale derived from the Denver Development Scales, and the POCL.

This battery represents a considerable change from Phase II – a result of Phase II* data analysis, experience of the SRI testers, and a set of local field tests completed by SRI during July 1976. Phase II/Phase III differences are as follows:

(1) Only the 32-item version of the Preschool Inventory will be retained. Analysis of Phase II data has shown that the loss in psychometric quality engendered by halving the PSI will be negligible. The rationale for using the 64-item PSI was the hope that a verbal subscale could be developed, but this has not been possible. Time saved in using the shorter version can better be used in measuring verbal and motor skills directly.

(2) Although the PSI is highly verbal in nature and correlates well with such measures as the Peabody Picture Vocabulary Test (PPVT), it will be helpful to have an explicit measure of verbal skills. The PPVT has been selected over the McCarthy Verbal Memory Scale (used in Phase II), since there are available norms established well beyond the target age and it is a "pure" language skill test, whereas the MVM is as much a memory test as a test of verbal skills. Furthermore, data collected during Phase II indicate a ceiling effect for the "list" items on the MVM.

(3) The Matching Familiar Figures Test (MFFT) and the Motor Inhibition Test have been dropped from the test battery. Phase II analysis indicates that the MFFT is not a valid measure of cognitive style in three- and four-year-old children, and this conclusion

---

* *See Chapter 3.*

85

is supported by recent literature.[3] The Motor Inhibition Test has not yielded acceptable short-term reliability, nor has a clear conceptual rationale for its use emerged from the literature.

(4) Finally, one major element of development in three- and four-year-old children is the emergence of a variety of motor skills. The NDCS test battery required a set of items selected from standardized tests that have face validity as an index of motor development in children enrolled in day care. Items from the Denver Development Scales were added because they contain fine and gross motor subscales useful for this purpose, are easy to administer, and require less than the 10-12 minutes available on Testing Day 2.

Phase III Child-Focus Observations

Child-focus observations were conducted in November 1976 and will be repeated in spring 1977. The Phase III instrument is a modified version of the Prescott observation system, virtually identical to the one used at the end of Phase II. Each target child is observed in one 20-minute period of free play during the morning. During the spring observations, each child will be observed by one black and one white observer. Balancing of observers by race was undertaken to control for possible race-of-observer effects. As noted in Chapter 3, Phase II data suggested that such effects might exist.

In order to further investigate possible observer effects, a check on inter-rater agreement will be conducted during the spring data collection. Pairs of observers, one black and one white, will observe the same child for one 20-minute sequence on the fifth and tenth day of the observation period.

In addition to indicating a possible need to control observer effects, Phase II findings led to another major modification in the Phase III child-focus data collection. As indicated in Chapter 3, many important child behavior variables did not emerge clearly in the course of naturalistic observation. Consequently, an effort has been made to enhance Phase III observation data by use of contrived play situations designed to highlight selected child behaviors toward peers and materials. The purpose of these is to raise the frequencies of behaviors that could not be adequately measured during Phase II, largely because the environment did not supply appropriate stimulus situations frequently enough.

As a result of a week-long field test conducted jointly by AAI and SRI, a stimulus situation was developed that appears to provide adequate levels of child interest, variability, and conceptual face validity. The situation provides an opportunity — but not the necessity — for cooperative play. Children are observed in same-sex pairs as closely matched in age as possible. The structured situation consists of two segments:

- In the *limited resources* situation, the children are given a Play-Doh game with one Play-Doh mold but an abundant quantity of Play-Doh. The crux of the situation is that only one child can use the mold.
- In the *abundant resources* setting, the children are given a Fisher-Price Play Family Village and associated materials. This toy permits independent play, cooperative play, and mutual fantasy play.

After introducing the children to the situation, the tester avoids initiating interaction with them as much as possible but responds appropriately to any questions or statements from the child. The tester counts the frequency of several elements of play and completes various summary codes as well as the POCL. Simultaneously, the other observer records time-sequence data as in any other Prescott observation.

Phase III Adult-Focus Observations

Adult-focus (caregiver) observations were conducted by SRI only once during Phase II of the NDCS but will be administered twice during Phase III. The first round of observations was completed in October 1976; a second round is being conducted in spring 1977. In each case, classrooms were or will be observed on two days but, during the spring observations, each classroom will be observed once by a white observer and once by a black observer. All observations are counterbalanced; white and black observers each do equal numbers of first observations. In classrooms where both teachers and aides are to be observed, black and white staff will observe teachers and aides equally often. The instrument used will be the modified SRI observation system administered in Phase II.

In addition, as in the child-focus observations, a check on inter-rater agreement will be conducted during spring data collection. Pairs of observers, one black and one white, will observe the same caregiver for one five-minute sequence on the fifth and tenth day of the observation period.

Finally, again in the spring, caregivers will be assessed using a new instrument based on the caregiver competences developed by the Child Development Associates Consortium. Caregivers will receive overall ratings in each competence area, and their performances will be checked against a list of specific behaviors related to each area.

Center Organization, Operations and Cost Data

Program and cost data were collected periodically throughout Phase II, and this effort is being continued in Phase III. Cost data are recorded by means of the Research Cost Accounting System (RCAS).* Information on enrollment, staffing and staff backgrounds, schedules, services offered, curriculum and other aspects of center operation are recorded by means of the Research Program Information System (RPIS).[4] While Phase III RPIS instruments have been slightly modified to provide more sensitive and accurate portraits of center organization and operations, RCAS procedures and forms are unchanged from Phase II.

Periodic Center Information

During the study year, at least two interviews will be completed by site staff with the center director and other center personnel to gather a more complete, in-depth portrait of the center and to complement other data collection efforts. The final interviews will be completed by mid-June 1977.

Parent Interview and Survey

In spring 1977, two instruments will be administered to samples of parents of target children. The first, a brief *Parent Survey*,** will be administered to approximately 400 parents chosen from all centers. The survey will yield data on parental attitudes and child-rearing practices; specifically, parents' aspirations for their children, parents' responses to some common situations, and children's reading and television habits. The second instrument, an in-depth interview, will be administered to 120 to 150 parents in 12 centers, four at each site. Data will be collected on parent preferences for different types of day care, parent goals for children in day care, parent involvement in day care and family use of center services. Results from centers funded entirely by the government will be compared to results from centers with mixed funding.

Day Care Forces Inventory

The Day Care Forces Inventory (DCFI) is an instrument designed to elicit caregivers' judgments about the importance and valence (positive or negative character) of various factors in the day care environment affecting their work — space, equipment, the director,

---

** See Chapter 4.*

*** Abt Associates is indebted to Virginia C. Shipman of Educational Testing Service for permission to use several of the questions devised for the ETS Head Start Longitudinal Study[5] and for her help in selecting the questions to be used.*

other caregivers and many additional factors. The instrument uses a sorting technique, in which caregivers place in order a set of cards listing features of the environment, first according to importance and then according to valence. It is being administered in spring 1977 to the 220 teachers and aides observed with the Adult-Focus instrument and also to the 57 directors of target centers.

Site Case Studies

The three sites of the NDCS – Atlanta, Detroit and Seattle – were selected to represent a wide range of geographic and demographic characteristics, as well as a diverse set of historical circumstances and administrative arrangements in day care. The success of this effort is shown indirectly by the large descriptive differences across sites which have emerged throughout Phase II data. In order to provide a qualitative context for the statistical analyses of site differences, case studies of day care will be undertaken at each site during the year. The case studies will provide, in narrative form, information about general economic, geographic and demographic characteristics of each site; the history of day care at each site; state and local regulations and regulatory bodies; the influence and local interpretation of federal regulation, particularly with regard to funding mechanisms; and the local day care market response to regulation.

Appendix A: NDCS MEASUREMENT AND ANALYSIS METHODOLOGY

The purpose of the National Day Care Study is to assess the costs and effects of alternative forms of center care, particularly of center characteristics that are currently or potentially subject to federal regulation. Phase II of NDCS was a naturalistic study. AAI analytic efforts in Phase II focused on developing a set of theoretically sound data-based constructs capturing the behavior of caregivers and children, and on devising a methodology to disentangle the complex relationships among center characteristics, center processes, and outcomes for children and parents. This appendix reviews a number of important measurement and analysis issues which were raised in the course of this effort.

Construct Development for Observation-Based Data

The development of constructs for the child- and adult-focus observation instruments has been discussed in some detail in Chapter 3. We confine discussion here to more technical considerations regarding these data. The NDCS is perhaps the first major national study which has placed primary emphasis in its data collection on direct observation of children in a group setting. While the reliability of the test battery in the NDCS (which during Phase III includes measures such as the Peabody Picture Vocabulary and the Preschool Inventory) is quite high, this is not true for the observation-based measures. In fact, defining an appropriate measurement framework for the observation data was itself a major issue. Inter-observer reliability is the usual focus of concern with observational instruments.[1] It is clear from our analyses of the Phase II data that the stability of the measures, and as a result their very meaning, is open to question. For example, while the child-focus observations were intended to tap enduring characteristics of children's behavior (e.g., task orientation and aggression), they may in fact be highly responsive to determinants of the immediate context (e.g., specific physical area within the center, type of activity period observed) and to characteristics of observers (e.g., sex, age, training, experience and race).

We have approached these concerns analytically through the Theory of Generalizability.[2] We can examine, with mixed and random effects ANOVA models, the relative contribution of the characteristics of those observed, the observers and the contexts to the total variability in the observational data. The Phase II data collection design permits only a partial analysis of these issues, since the original assignment of observers to centers, classrooms, children and caregivers was not specifically planned with this in mind. As a result of the limited Phase II analyses, however, the NDCS elevated the importance of this investigation. Modifications were instituted in the Phase III data collection procedures to permit a fuller variance components analysis of the data. The results of these psychometric analyses are of critical importance in directing the appropriate use of the data in subsequent policy analyses. The interested reader is referred to the *Phase II Research Report* for more details.

Analysis Plan Development for NDCS Phase II

Since Phase II of the NDCS is a naturalistic study, a highly sensitive quantitative methodology is necessary to isolate the effects of policy regulatable variables and their interactions from the many other center and child factors confounded with them. In particular, Chapter 1 of this report noted that there were differences among the various types of day care centers in the entering characteristics of children and their families. This non-equivalent grouping of children by centers has resulted from different center funding mechanisms, different child enrollment procedures, self-selection effects and demographic differences across sites. Since children are not initially the same from center to center, this may introduce differences in center processes and subsequent rates of development, even in the absence of differences among centers on the policy variables. Thus, apparent differences in the effects of centers may not be the result of observed differences in the policy variables, but rather be attributable to other factors on which the centers may also vary. The 49-center study and eight-center study in Phase III were specifically designed to minimize the potential confounding influences of these other explanatory variables. In the absence of such design controls in Phase II, we must rely heavily upon statistical methods to adjust out the effects of these other factors.

The General Linear Model is perhaps the most widely used methodology for this purpose. Extensive discussions of the applications of these methods to social and educational program research can be found in Kerlinger and Pedhazur[3] and Cohen and Cohen,[4] with a more technical presentation available in Searle.[5] This method involves the formulation of analysis models which incorporate the independent variables of primary interest (i.e., the NDCS policy variables and interactions) and other factors confounded with these variables (i.e., the child and center variables, which are usually referred to as covariates). In addition, this approach allows us to consider in our analysis both continuous (e.g., a child's pretest score) and categorical (e.g., a child's sex) independent variables, and to examine both general and possible interaction effects associated with these variables.

In terms of the NDCS policy variables (staff/child ratio, group size, staff experience and education), we have two options for representing these variables within the General Linear Model. We could categorize each of these variables and, in effect, create a four-way crossed analysis of variance design. Alternatively, we can utilize them as continuous measures, as we might in a regression analysis. In Phase I of the NDCS, each center was in fact categorized on each dimension during the process of selecting a representative center sample (cf. *First Annual Report* for details). Available survey data on the policy variables was used for this purpose. During Phase II and III, however, very precise measurement of these variables was undertaken. Categorizing the independent variables in such cases is tantamount to throwing away valuable information about the variability among centers within any cell of the design. Thus, in the interests of maximizing the power of the NDCS analyses, we seek to capture this variability and will generally treat the policy variables as continuous independent measures.

In addition to the policy variables and their interactions, our General Linear Model analyses of this data must include a set of covariates which represent the other factors confounded with the policy variables. In selecting this covariate set, two principles suggested by Cronbach[6] have been followed: (1) the potential covariate must be significantly related to the outcome or dependent variable of interest in a given analysis; and (2) there must be substantial confounding among the NDCS centers on this dimension. In principle, this analytic approach results in a reduction in the bias associated with the initial non-equivalence of groups, and a gain in the precision with which we can estimate the effects of the policy variables.

Analysis of Process Data. In Phase II, the Prescott Observation Procedures were field-tested and modified for the specific purposes of the NDCS. As a result, there were substantial changes in the coding procedures from Fall 1975 (T1) to Spring 1976 (T2) which negate the possibility of meaningful comparative analyses across the two time points. Thus, the Phase II analyses of these data were limited to separate applications of the General Linear Model at each observation point.

Pending the results of the generalizability study being conducted in Phase III, the psychometric properties of the observation-based constructs may be adequate to support comparison across Spring 1976 (T2), Fall 1976 (T3) and Spring 1977 (T4). Several different types of analyses are possible here. The policy variables could be examined against profile averages across T2, T3 and T4 for each center or classroom. Such analyses would focus on possible level differences across centers in classroom processes which are attributable to differences across these units in the policy variables. Alternatively, we could investigate potential changes over time in classroom processes resulting from the additional resources provided for some NDCS centers as part of the Phase III design. Whether these analyses prove feasible or not, separate examination of the T3 and T4 data will be conducted to parallel the analyses already completed at T1 and T2.

In addition to examining theoretical constructs developed from the observational procedures, the NDCS intends to examine important individual codes within the Prescott instrument. Many of these individual codes represent potentially critical events for a child (e.g., crying, being disciplined), but do not occur with great frequencies. Although there may be some nonzero probability of such an event occurring with any individual child (or caregiver), no single occurrence of this event may be recorded for this person during the observation period. From a methodological perspective, these codes are best conceived as dichotomous variables which indicate simply incidence or nonincidence of the event. With a dichotomous outcome variable, particularly with low incidence rates (e.g., less than 10% of the cases with a positive response), use of the General Linear Model results in inefficient and potentially biased estimation[7] of effects. A more appropriate analytic tool in this case is the Logit Model. In form it is identical with the General Linear Model in that the policy variables, their interaction terms and a covariate set are included as independent measures, but estimation of effects is achieved through the method of maximum likelihood under the assumption that the incidence variables are generated in accord with a binomial distribution. Hanuschek and Jackson[8] describe this procedure in more detail.

Analysis of changes in cognitive test scores. For well over a decade[9] there has been considerable discussion among methodologists about the appropriateness of alternative methods for analyzing change.[10] Much of the recent discussion has focused on alternative suggestions for the appropriate use of pretest information on the non-equivalent groups in naturalistic or quasi-experimental studies. Most methods suggested share a common form. In a simple treatment/control group study, this form can be represented as:

$$\alpha = \bar{Y}_p(t_2) - \bar{Y}_c(t_2) - \beta[\bar{Y}_p(t_1) - \bar{Y}_c(t_1)]$$

where α = effect estimate of the treatment or policy variable;

$\bar{Y}_p(t_2)$; $[\bar{Y}_c(t_2)]$ = posttest mean for program [control] group; and

$\bar{Y}_p(t_1)$; $[\bar{Y}_c(t_1)]$ = pretest mean for program [control] group.

They differ in terms of their approach to estimating the β coefficient which is used to adjust for pretest mean differences among the groups in any design. In terms of adjustment for other covariates, such as social class variables, the methods tend to revert to a common General Linear Model approach.

If there is any consensus in this debate, it is that no one method can be applied with uniform success in all cases. As Kenny[11] suggests, the best that we can hope to achieve is the selection of an analyses model which is appropriate, conditional upon some assumptions about the manner in which the data were generated. With this caveat in mind, the NDCS has adopted a multiple analysis methods approach. As Cook[12] has argued, when no one method is clearly best, the use of multiple analysis strategies, each with its own strengths and weaknesses, can substantially enhance our confidence in the conclusion that we ultimately draw about the effects of the policy variables.

In selecting specific strategies for NDCS analyses of the cognitive battery, the major problem we faced was that young children are developing and change over time on cognitive dimensions, even in the absence of effects attributable to interventions such as day care. As Bryk and Weisberg[13] have shown, however, the application of traditional adjustment methods, such as simple gain scores and analyses of covariance, can result in these cases in a gross distortion in our estimates of program effects. Since children are not initially the same from center to center, we expect that they will develop at different rates even in the absence of effects associated with differences between centers on the policy variables. We have developed in the course of the Phase II analyses three alternative statistical approaches, each of which attempts to sort out differences in effects associated with the policy variables. These three methods are:

(1) residual gains analysis – for each child (or grouping of children) a T2 score adjusted for natural growth is computed of the form:

$$Y_i(t_2)_{\text{adjusted}} = Y_i(t_2) - \beta\,[\,Y_i(t_1)+C]$$

where $i = 1 \ldots M$ denotes individual subjects;

β is an adjustment coefficient determined from extensive exploratory analysis of the pretext/posttest data; and

C is a fixed increment so that $Y(t_2)_{\text{adjusted}}$ (the mean of adjusted scores) equals $Y(t_2)$, where $Y(t_2)$ is the actual posttest grand mean.

These adjusted posttest scores are then subject to a General Linear Model analysis as described in the previous section.

(2) percentage maximum possible gain – this method[14] consists of changing the tests score metric from that of a simple or adjusted gain as used above to an expression of the observed gain as a function of the maximum possible gain, i.e.,

$$Y_i(t_2)_{\text{transformed}} = \frac{Y_i(t_2) - Y_i(t_1)}{Y_{\max} - Y_i(t_1)}$$

where $Y_{\max}$ is the maximum possible test score, e.g., the total number of items on the test; and

$Y_i(t_2)_{\text{transformed}}$ represents the percentage of items unknown at T1 which is learned between T1 and T2.

These scores in the transformed metric are then subject to a General Linear Model analysis.

(3) value-added analysis – this technique[15] explicitly attempts to predict a posttest status for any group of children exposed to some day care condition, as if they had been subject to some natural or control condition. Available covariate information can be used in the prediction process. An estimate of the effect of a specific day care environment, beyond the anticipated natural development, can be formulated as:

$$\hat{\alpha}_j = \overline{Y}_j(t_2) - \overline{Y}_j(t_1) - \pi_j\,[\bar{a}_j(t_2) - \bar{a}_j(t_1)]$$

where $\hat{\alpha}_j$ is the estimated effect of day care environment j;

$\bar{a}_j(t_1)$ and $\bar{a}(t_2)$ are the average ages of the children in group j at T1 and T2 respectively; and

π_j is an estimate of the average growth rate for subjects in group j for the time period from T1 and T2.

A discussion by Goodrich[16] provides a more extensive rationale for these methods. A comprehensive discussion of the application of these methods to the NDCS will be contained in the final research report.

Sensitivity Analyses. Despite the comprehensiveness of the analytic plan outlined above, there are several other factors which could potentially threaten the validity of our results. These concerns include: missing data on individual variables for some cases; subject attrition from the study before completion of T2 (or T4); possible sensitivity of our estimated effects to the presence of outliers and fringeliers[17] in the data set; and errors of measurements in the independent variables. While serious problems in any one of these domains are not anticipated, nevertheless the NDCS plans, in Phase III, to undertake a series of validation studies where each of these potential concerns is considered, and the sensitivity of the estimated policy effects to alternative assumptions in each domain is assessed.

Unit of Analysis Issue

The choice of a proper unit of analysis for the NDCS represents a difficult analytic and conceptual challenge. This is a complex problem, particularly in education contexts which routinely consist of hierarchical systems. For example, the NDCS design consists of children nested within classrooms, classrooms within centers and centers within sites. There is a growing literature on the problems associated with the analysis of such data with Haney[18] and Cronbach[19] representing excellent reviews.

For the NDCS, there are three possible units of analysis: the child, the class and the center. For any change analysis the class level is problematic. This organizational structure is not a stable entity. It often undergoes a complete alteration in the course of just a few months. For the single time point analysis of the observational data, however, a classroom unit is feasible and it constitutes, perhaps, the most conceptually satisfying level for an examination of the interactional processes which define the day care context.

For any change analysis, such as those described for the cognitive test battery, the child and the center represent the only feasible alternatives. Cronbach[20] musters impressive statistical arguments for initiating analysis at the most aggregated level – in this case the center. In addition, since the policy variables can only be defined for an aggregate, such as classroom or center, the center comes closer to a sufficient definition of the unit of treatment. On the other side, however, the effects of the policy variables are actually experienced at the individual child level. Further, since there are substantial differences in the background characteristics of children attending the various NDCS centers, Haney's work suggests that adequate adjustment for these confounding influences can only be obtained by analysis at the child level.

Since again we see no obvious resolution of this conflict, the NDCS will analyze its data at both the center and child level. A multilevel analysis as suggested by Cronbach[21] will be undertaken. If the results of analyses from different levels do not agree, we will use aggregation theory[22] in an attempt to sort out the likely determinants of this conflict.

Power Investigation of the Phase II Natural Experiment

Even under the assumption that the analysis plan outlined above is highly effective in minimizing the biasing effects associated with pre-existing differences among the NDCS centers, the power of the naturalistic experiment to detect meaningful policy effects is an important issue. In particular, for any center-level analyses where we have 64 units within three sites, a sample of this size is not sufficiently large to automatically guarantee the visibility of meaningful relationships between the policy variables and measured classroom processes and child outcomes. Thus, a statistical power analysis was conducted in Phase II to investigate the likelihood of detecting effects in our center-level analyses, if in fact they were present.

Strictly speaking, power is the probability that a given statistical significance test results in a rejection of the null hypotheses. In general, the power of a particular test depends upon the sample size, the variability in the outcome measure, the size of the real effects, and the level of significance adopted. In addition, power is conditional upon fallibility of the dependent and independent variables and upon the degree of multicolinearity among the independent variables. These latter factors make direct computation of power difficult. For this reason, the Monte Carlo method was employed. Simulations were conducted based upon a range of assumptions concerning the true magnitude of effects on cognitive instruments such as the Preschool Inventory. For purposes of the Monte Carlo study the following conventions for describing effect sizes were adopted:

Effect Size	Number of points gained on the PSI	% of Variance Explained (R^2)
Very small	½	.02
Small	1	.06
Medium	1½	.14
Large	2	.25

Thus, for example, we would say that one of the policy variables (e.g., staff/child ratio) had a "small" effect if children exposed to high ratios typically gained one point more on the PSI from fall to spring than did children exposed to low ratios. (Normal growth is about one point per month.) An effect of this magnitude would imply that ratio experience accounted for six percent of the variance in PSI gains from fall to spring.

The results of these power analyses were as follows:

- for main effects of any of the four variables, power is marginal for small effects, adequate for medium size effects, and very high for large effects. These results indicate that a null finding for a main effects, based on all centers in the three sites, would imply that the real effect is certainly not large and probably not medium size. Only small effects are likely to escape detection.
- the power to detect interactions is high where interactions have about the same magnitude as main effects. For small interactions, the power is not sufficient, but with large interactions, we are practically certain to detect these effects in a careful analysis.
- there is limited power at the center level for any within site analyses. This means that practically significant effects would not be detectable unless we relax the statistical significance criterion to .10 or higher.

It should be emphasized that while a number of alternative cases were considered in the course of these center-level power studies, the results reported above are conditional upon a set of relatively simple analytic models for the PSI measure. One must be cautious in attempting to draw inference beyond the boundaries of the Monte Carlo study. These results do not necessarily apply to other measures in NDCS, such as the observational constructs, or to alternative analyses on the PSI.

Utility of the NDCS Findings

Our ability to generate policy relevant conclusions from the NDCS findings depend upon two issues: the internal validity of our analyses, and the representativeness of the 64 sample centers and three sample sites of the larger day care universe. The general design and analysis plan for the NDCS is a very purposeful attempt to safeguard the study from obvious threats to its internal validity. Thus, the major question to be considered here is one of population validity[23] or "what kinds of inferences are reasonable to draw from the NDCS sample to the day care universe?"

The most important question which this study seeks to answer is "are there *general* relationships that exist across all day care contexts between regulatable characteristics of centers, and child process and development variables?" Thus, the site and center selection plan for Phase II and Phase III was developed so as to subject this global hypothesis to as general and severe a test as possible within limits of reasonable study cost. To accomplish this, the NDCS deliberately selected sites and centers to maximize the diversity within the sample. In so doing, we maximized the power of our analyses to detect nonhomogeneous effects across different day care contexts.*

Site Effects. In terms of generalizing beyond the three sites included in the study, the most easily interpretable finding would be that the effects of the major study variables are essentially the same in each study site. Given the diversity across sites, we would then feel confident that the study findings can be generalized to the day care universe.

If the pattern of results is more complex, however, involving different results across sites, the process of attempting to draw general recommendations from this study to the larger federal scene is more complex. It should be noted that the presence of apparent site effects in the data may easily result from the heterogeneous populations across sites. Thus, a careful analysis is necessary to separate effects due to diversity within the NDCS sample from those attributable to other more policy-relevant characteristics of sites (e.g., local regulations and funding mechanisms). The *First Annual Report* describes a prototype procedure which may be useful in determining whether national generalizations are possible, and what the limits on these generalizations might be. This procedure will be used in Phase III analysis should it be determined necessary.

It should be pointed out that a random selection of sites and centers for the study would not have been sufficient since it is unlikely that a random selection of three sites and 64 centers – constraints determined by the cost of the study – would have resulted in the maximal diversity among units necessary to assure optimal conduct of the study. Further, a sample of only three sites is too small for random selection and probability theory to be useful; the value of random sampling can only be obtained when the number of units selected is much larger.

Appendix B: NATIONAL SUPPLY STUDY

In April 1976, the NDCS began the National Day Care Center Supply Study, a survey of the characteristics of center-based day care and its users, based on a telephone survey of 3,000 randomly selected day care centers throughout the United States. The Supply Study was undertaken to provide a detailed and current statistical profile of center care in the U.S. that would assist the interpretation and generalization of NDCS findings and could be used independently by national and state day care planners and operators. The Supply Study design is separate from Phases II and III of the NDCS; a first round of data collection activities for the Supply Study, however, included the 64 NDCS centers in order to provide a connection between the Supply Study and the NDCS data bases.

Currently available data on child care in the U.S. have serious limitations. For the most part these data are either several years old or narrow in scope. There are several questions that need to be addressed and some important data that need to be collected before re-evaluation of federal day care policy can be completed. There is a need for up-to-date and nationwide information on the characteristics of the programs, staff and physical facilities of licensed day care centers. For example, the data provided by the Westinghouse/Westat study in 1971 is out of date, and its sample of day care centers is too small for state-by-state analysis or extensive econometric modeling. The 64 centers in the National Day Care Center Study also comprise too small a sample from which to extrapolate the characteristics of the national population of centers. Because of the increasingly important role state governments play in the determination of child care policies, data on the supply of center-based care should be available on a state-by-state basis.

The Supply Study has the objective of providing:

- a statistical profile of the characteristics of day care centers in each state and an aggregate profile at the national level;
- estimates of the degree to which centers are in compliance with the key provisions of state and federal day care regulations;
- an econometric model of the responses of day care centers to variations in federal day care regulation, local monitoring practices and Title XX reimbursement rates; and
- a data base that can be used to extrapolate the national implications of the cost-effectiveness findings of the NDCS.

The statistical profile will describe the size of the industry; characterize the children, client-families and staff who comprise center day care; identify the relative frequency of different program characteristics and financial arrangements; and examine the degree of ethnic and socioeconomic integration in center care.

Estimates will be developed of the degree to which centers are in compliance with several key provisions of state and federal day care regulations. The following are the most important areas in which the degree of compliance will be examined:

- *Over-Enrollment:* the difference between enrollment and licensed capacity;
- *Staff/Child Ratios:* the difference between the number of staff a center should have, given the size and age distribution of its enrollment, and the number of staff actually available;
- *Group Size:* the difference between the maximum number of children allowed in a group, given the age of the youngest child in the group, and the actual group size; and
- *Parent Participation:* the percentage of centers in compliance with requirements

that Parent Advisory Councils have a role in the center's decision-making.

An econometric model of the day care market will be developed and estimated from the Supply Study data base. The model will be designed to forecast the impact of alternative licensing regulations, alternative monitoring and funding strategies, and variations in local economic conditions on the following set of variables:

- licensed capacity
- enrollment
- number of paid staff and average wage
- average cost of child care per child-year
- average parent fee
- total private-paid enrollment
- numbers of profit and non-profit centers

When the model is complete, projections will be made of each of these variables for each version of the FIDCR and for alternate reimbursement policies.

One of the most important motivations for the Supply Study was the assistance it could provide in projecting the cost-effectiveness findings beyond the 64 centers participating in the NDCS. The data available from the telephone survey will identify the major types of day care centers which exist nationally and the degree to which each type is represented among the 64 NDCS centers. In addition, the national survey will serve to extrapolate the national cost implications and the potential impact on the day care market of the final recommendations of the cost-effectiveness study.

The data collected by the Supply Study will come from a probability sample of approximately 3000 day care centers. The sampling unit is a day care center which:

- provides nonlive-in child care services;
- has a capacity for 13 or more children;
- is open at least 25 hours a week for at least 9 months a year;
- does not specialize in the provision of care to handicapped children.

These sampling criteria differ from those used to select the Phase II sample in the following ways:

- The Supply Study excluded centers with a licensed capacity for less than 13 children (the FIDCR definition of centers as opposed to group day care homes) while the NDCS excluded centers with enrollments of less than 25.
- The NDCS excluded centers with fewer than 20 *full-time* three- and four-year-olds. The Supply Study excluded centers that enrolled only part-time children; age distribution was not a selection criterion in the Supply Study.
- The NDCS focused primarily on centers currently serving or eligible to serve publicly subsidized children. The NDCS also attempted to include centers that were racially and economically integrated. No such criteria were used in the Supply Study center selection.

As a consequence of these differences in sampling criteria, the Phase II centers differed from the Supply Study's national probability sample in several important respects, illustrated in Table B-1. Centers in the Supply Study sample had smaller enrollments and fewer staff than the Phase II centers. They served fewer black families and fewer low-income families. More of them were operated for profit than was the case in the Phase II sample. However, despite the fact that age of children served was not a selection criterion in the Supply Study, the percentages of three- and four-year-olds in the two samples were virtually identical.

Stratified random sampling was used to select centers for the Supply Study telephone survey. Day care centers were stratified by the state in which they operated, and a random sample was drawn from each stratum. The sample size in most states was proportional to the states' population of centers, but the sample size was increased in states with few centers to provide a mini-

Table B-1

Comparison of Supply Study Proportional Sample and Phase II Centers

	Supply Study Proportional Sample[a]	Phase II Centers[b]
Average Licensed Capacity	55	77
Average Enrollment	54	72
percentage full-time	74.1	83.3
percentage part-time	25.9	16.7
average full-time equivalent enrollment	47	65
percentage white	65.8	45.2
percentage black	27.0	50.6
percentage other minorities	7.2	4.3
percentage less than three years old	13.4	16.2
percentage three- and four-year-olds	49.4	49.9
percentage five and older	37.2	33.9
Average Number of Families Served	48	62
percentage with income less than $6000	30.3	39.3
Legal Status (Auspices) of Centers		
percentage organized for-profit	40.4	23.4
percentage nonprofit	59.6	76.6
Average Number of Staff	11	17
percentage full-time paid	52.8	50.3
percentage part-time paid	29.6	29.3
percentage volunteer	17.6	20.4

[a] Supply Study data are based on the first rounds of the telephone survey, conducted during April-May 1976. The data are therefore preliminary and may change somewhat when information from all four rounds becomes available.

[b] To insure comparability of data, information on the Phase II centers shown in this table was drawn from telephone interviews with directors of the 64 centers conducted during the first round of the Supply Study. Also, data on children and families cover full center enrollments, not just target children. For these reasons, percentages shown in the table differ from those in the body of the report, which were based on extensive on-site examination of the centers and interviews with parents of target children.

mum level of accuracy in all state-level profiles of center characteristics.

The telephone survey was conducted in four quarterly rounds of 750 interviews each. The samples for the four rounds were drawn independently, and the telephone interviews in each round were completed within a four week interval. The spacing of the 3000 interviews over a twelve-month period permits estimation of seasonal variations in center characteristics.

Another sample consisting of about 70 centers participating in the telephone survey will be used for an on-site survey. The on-site survey will serve to assess the accuracy of responses to the telephone survey; to collect detailed and documented information concerning staffing and grouping patterns, services provided to children and families, and financial status of centers; and to measure certain key center characteristics in precisely the same way as they are being measured in the centers participating in the NDCS.

A report of the findings of the Supply Study will be available in November 1977.

Appendix C: INFANT CARE SURVEY

The Infant Care Survey will provide descriptive information on both a national and regional basis about the characteristics of center-based day care for children from birth to three-years-old. This study has been commissioned by the Office of Child Development, Department of Health, Education and Welfare, as a separate task within the National Day Care Study. The major focus of the NDCS is upon three- and four-year-old children. It is widely believed, however, that the use of day care centers for children under three is increasing in the United States. The Infant Care Survey will supplement the information gathered by the NDCS and will partially fill the need for more data about the types and availability of center care for younger children.

Recent research reports and ongoing research efforts[1] reflect a national interest in the effects of group care upon the development of very young children. These studies have tended to focus on differences between family and group care rather than on effects of environmental variation within group care, such as differences in staff/child ratio, caregiver training or center philosophy. The critical variables that are assumed to distinguish family from group care are the separation of the child from his or her mother and association with a relatively large, unrelated peer group. However, detrimental effects of group care have not generally been found. In fact, for children from disadvantaged backgrounds, there is some evidence that center care in intensive intervention environments leads to cognitive gains, and where the parents are involved, differences in parent/child interaction are found.[2] These recent studies are in marked contrast to older studies of institutionalized infants, which consistently showed negative effects of group care relative to the home environment.[3] Clearly, differences in the quality of group care have had a crucial influence. Unfortunately, the factors contributing to quality have not been systematically identified, although a large number of highly focused laboratory and field studies have provided relevant information.

State and federal regulations have been developed in the absence of definitive research findings. Generally, expert opinion, cost factors and political considerations have been the determining factors in the development of day care standards for young children. The Infant Care Survey is an attempt to provide policymakers with three kinds of data that are not currently available. First, a telephone survey (an extension of the NDCS Supply Study) will provide data on the national and regional distribution of infant care. Second, on-site interviews will provide data regarding the critical issues and variables which, from the viewpoint of providers and consumers of center-based infant and toddler care, relate to quality of care. Third, direct observations of caregivers will provide descriptive information on center processes in infant care.

The Telephone Survey

The telephone survey will consist of four rounds of interviews, conducted as part of the Supply Study. Three types of centers will be examined.

• Infant:	Centers with five or more children under two years old.
• Toddler:	Centers with fewer than five children under two but more than five children under three years of age.
• Preschool:	Centers with fewer than five children under three years of age.

The survey data will be used to compare infant and toddler care centers with predominately preschool centers. Means, medians, ranges, frequency counts and other descriptive statistics will be calculated for center-wide characteristics. Center size, recent growth, family demographics, center fees, costs, services, staff characteristics and center organization will be examined in order to determine similarities and differences in center types. Breakdowns of the characteristics of the three center types by state and region will be computed.

Additional analyses of infant center characteristics will focus in more detail on care for infants and toddlers. Education, age, training and experience of caregivers will be examined by age of child, as will cost and fee structures and classroom age-mixing policies. The study also will examine staff turnover in infant and toddler classrooms. Regional breakdowns of within-center patterns will be examined to determine how the quite different day care regulations and child care traditions in different regions have affected the development of center-based care.

Finally, data will be analyzed to identify seasonal or general trends in enrollment and staffing patterns, classroom structures and characteristics of families using day care. Regions will be separated in the analysis in order to determine whether trends are localized or national.

On-Site Interviews and Observations

The sample for the on-site interviews and observations will be a cluster of centers in several cities selected on the basis of geographical diversity. Centers with different levels of staff/child ratio, staff education and staff experience will be selected. The data base developed from the center visits will provide policy-makers with more in-depth information about the concerns of consumers and providers of infant/toddler care and help to identify factors in the day care environment which might be regulated to the benefit of children and families.

The issues to be developed in interviews and observations are somewhat distinct. Interviews are useful in exploring broad and subjective opinions and concerns. Observations will be focused upon detailed and specific environmental events and will be used to compare experiences for children when characteristics of care are varied.

Interviews will be conducted with directors, staff and parents. Analysis of the interview data will concentrate upon expanding the base of information about how caregiving staff and parents feel about the issues involved in day care for very young children. These issues include the role of regulations, the role of parents, satisfaction of staff with job conditions, and child-rearing attitudes as they relate to center care for children under three years old. Analysis of the observation data will focus upon the relationships among the characteristics of the caregiving situation (e.g., staff education and experience, staff/child ratio, age of children in care) and the behavior of the caregiver (e.g., classroom activity pattern, caregiver involvement, stability of caregiver behavior).

Appendix D: THE NDCS CONSULTANT PANEL, OCD STAFF AND CONSULTANTS, AND NDCS STAFF

The Office of Child Development has recognized from the study's inception the need to subject the design, analyses and findings of the National Day Care Study to rigorous scrutiny by technical and policy experts. Abt Associates Inc. organized a panel of consultants which has met periodically throughout the study. This nationally selected group of experts has reviewed all major reports, advised the NDCS staff and made important suggestions that have improved the design and implementation of the study. A Black Task Force subgroup was organized during Phase II of the study. Its members, drawn from the advisory panels of the NDCS and the companion National Family Day Care Home Study, have identified technical and policy issues of particular concern to minorities, especially in reviewing the NDCS outcome measures as to their appropriateness and significance for minority children. In addition, they have fostered the increased involvement of black scholars and day care specialists in the NDCS research.

NDCS CONSULTANT PANEL

The members of the NDCS consultant panel were all selected on the basis of long-standing interest and expertise in day care and early childhood programs.

Robert O. Bland, Ph.D.	Assistant Vice President, Merrill-Palmer Institute, Detroit, Michigan
R. Virginia Burke	Director, NIH Day Care Center, Rockville, Maryland
Jean V. Carew, Ph.D.	Visiting Scholar, Boys' Town Center for the Study of Youth Development, Stanford University
John R. Dill, Ph.D.	Assistant Professor of Psychology, City College of the City University of New York, and Research Director, Day Care Coordinating Council of Essex County, New Jersey
Arthur C. Emlen, Ph.D.	Director, Regional Research Institute for Human Services, Portland State University
Robert Allen Fein, Ph.D.	Psychotherapist, Evaluator, McLean Hospital Lecturer on Androgyny and Family Policy, Massachusetts Institute of Technology
Asa Grant Hilliard III, Ed.D.	Dean, School of Education, San Francisco State University
Mary C. Howell, M.D.	Medical Director and Pediatrician, Charles Drew Family Life Center, Dorchester, Massachusetts
Teh-wei Hu, Ph.D.	Professor of Economics, Pennsylvania State University
James Alan Levine	Director, Child Study Center, Wellesley College, and Ford Foundation Travel-Study Fellow
Richard J. Light, Ph.D.	Professor of Education, Harvard Graduate School of Education
John M. Love, Ph.D.	Director, Research Department, High/Scope Educational Research Foundation, Ypsilanti, Michigan
Christine Marston	Coordinator, King County Child Care Coordinating Committee, Seattle, Washington
Keith McClellan	Director, Perlis Remotivation Center, Cleveland, Ohio

William J. Meyer, Ph.D.	Professor of Psychology, Director of Developmental Psychology Training Program, Syracuse University
Gwen G. Morgan	Lecturer, Lesley College: Problems and Issues in Day Care; Consultant, Texas Office of Child Development, Southeast Educational Conference
John H. Niemeyer	President Emeritus, Bank Street College of Education; Consultant on Child Care
Daniel M. Ogilvie, Ph.D.	Associate Professor of Psychology and Chairman of Graduate Program, Livingstone College, Rutgers University
Elizabeth Prescott	Director of Research, Pacific Oaks College
Mary Potter Rowe, Ph.D.	Special Assistant to the President and Chancellor for Women and Work, Massachusetts Institute of Technology; Specialist, Economics of Day Care
Richard Roy Rowe, Ph.D.	Vice President and Director of the Cambridge office of American Institute for Research (AIR); President, Massachusetts Children's Lobby; Lecturer, Harvard Graduate School of Education
Nancy Ewing Travis	Director, Child Care Training Project, Southern Regional Educational Board, Altanta, Georgia
Harmon L. Walker	Director, Seattle-King County Head Start Program, Seattle, Washington
James Clayton Young, Ed.D.	Associate Professor of Early Childhood Education, Georgia State University

OFFICE OF CHILD DEVELOPMENT

Preston Bruce, Jr., Ph.D.	Chief, Day Care Services Division, Office of Child Development, HEW
Allen N. Smith	Project Officer, National Day Care Study, Day Care Services Division
Carol Spence, Ph.D.	Assistant to the Project Officer, National Day Care Study, Day Care Services Division

OFFICE OF CHILD DEVELOPMENT/ CONSULTANTS

Robert L. Brennan, Ed.D.	Senior Research Psychologist, American College Testing, Development Research Department, Iowa City, Iowa
Victor G. Cicirelli, Ph.D.	Professor of Developmental Psychology, Purdue University
Marvin G. Cline, Ph.D.	Professor of Educational Research, College of Education, Virginia Polytechnic Institute

ABT ASSOCIATES STUDY STAFF

William Bache III, Ph.D.	Senior Analyst
Christine Bornas	Study Secretary
Katharine Burleson	Programmer
Daniel Calore	Senior Analyst/Supply Study
Craig Coelen, Ph.D.	Director/Supply Study
David Connell, Ph.D.	Deputy Project Director
Myrtice Dye	Data Coordinator/Atlanta
Steven Fosburg	Senior Analyst
Naomi Fujimoto	Site Coordinator/Seattle
Mary Helen Gardner	Data Analyst/RPIS
Frederic Glantz, Ph.D.	Director/Cost Study
James Gokhale	Programmer
Nancy Goodrich	Director/Information Management System
Robert Goodrich	Director of Research

Barbara Goodson, Ph.D.	Senior Analyst
Marjorie Gorman	Analyst/RCAS
Janet Grasso	Data Analyst/RCAS
Michelle Green	Analyst
Muriel Hamilton	Site Coordinator/Atlanta
Margaret Hansen	Data Coordinator/Seattle
John Harrington	Data Coordinator/Atlanta
Carolyn Hawkins	Site Coordinator/Detroit
Kathryn Hewett	Senior Analyst
Jean Layzer	Director/Parent Study
Gayle Leach	Data Coordinator/Detroit
Arthur Mathis, Ph.D.	Senior Analyst
James McIntosh	Programmer
Julia Michaels	Analyst/Supply Study
Ricardo Millett, Ph.D.	Director/Site Case Studies
Richard Ruopp, Ed.D.	Project Director
Janet Smith	Data Analyst/RPIS
Jann Swanson	Analyst/Supply Study
Barbara Tannen	Analyst/Supply Study
Jeffrey Travers, Ph.D.	Associate Project Director
Donna Warner	Coordinator of Data Systems; Director/RPIS

NOTES TO THE CHAPTERS

INTRODUCTION

1. Committee on Finance, United States Senate, *Child Care: Data, and Materials,* Table 29 (Washington, D.C.: U.S. Government Printing Office, 1977, Draft, in press).

2. NDCS *First Annual Report, 1974-1975,* Volumes I and II (Cambridge, MA: Abt Associates Inc., 1976).

3. *ibid.*

4. Stallings, J., Wilcox, M. and Travers, J., *Phase II Instruments for the National Day Care Cost-Effects Study: Instrument Selection and Field Testing* (Menlo Park, CA: Stanford Research Institute, 1976).

5. NDCS *Phase II Research Report* (Cambridge, MA: Abt Associates Inc., 1977). Available on request from the Day Care Services Division, Office of Child Development, 400 6th Street, S.W., Washington, D.C. 20013.

Chapter I: PHASE II DESIGN AND SAMPLE

1. A detailed discussion of the screening criteria used to select the Phase II study centers appears in the NDCS *First Annual Report,* Volume I, *op. cit.,* pp. 12-13, 44-45.

2. To assess the significance of the screening criteria used in choosing the 64 study centers, data from the Supply Study's national probability sample were used to determine the percentage of centers which would have met the NDCS criteria. The only criterion that was at all ambiguous was the restriction that centers be located in or near a major urban area. In the analysis of the Supply Study data, "urban area" was defined as any of the 243 Standard Metropolitan Statistical Areas used in the 1970 Census. Such areas accounted for 68.6% of the U.S. population in 1970. (U.S. Bureau of the Census, *Statistical Abstract of the United States: 1974,* Washington, D.C., 1974).

3. The number of target classes fluctuated over the course of Phase II, as new children enrolled and others dropped out, and as classes were reorganized. The largest number at any one time was about 150.

4. It was necessary to include in the research sample a significant number of centers which relied heavily on parent fees as a source of income. The inclusion of these centers added diversity to the sample on a number of dimensions – the socioeconomic status of families, staff/child ratio, group size and caregiver qualifications. It should be noted, however, that many of the "parent-fee" centers enrolled some FFP children. While parent fees were the largest income source for 33 centers, income from parents was the sole income source for only 12 centers.

5. U.S. Congress, Senate, Committee on Finance, *1977 Report on Child Care,* 93d Congress, 2d Session, 1977, Table 43.

6. *Federal Interagency Day Care Requirements,* Sept. 23, 1968 (Reprinted 1970). Section III (2), p. 9.

7. In more technical terms, if any policy variable accounts for 25 percent of the variance in any process or outcome measure, the probability of detecting such an effect approaches 1.0 at a significance level of .10. If a policy variable accounts for 14 percent of the

variance in a process or outcome measure, the probability of detection is .8 at a significance level of .10. If a policy variable accounts for only 6 percent of the variance, the probability of detection falls to .6.

Chapter 2: FIDCR POLICY VARIABLES

1. NDCS *First Annual Report,* Vol. I, *op. cit.* pp. 2-5.

2. A thoughtful discussion of the role of staff/child ratio in facilitating a wide variety of developmental experiences can be found in Fowler, W., "How Adult/Child Ratios Influence Infant Development," *Interchange,* Vol. 6, No. 1, 1975, pp. 17-31. Fowler is primarily concerned with very young children but many of his points apply to preschoolers as well. Another recent review of the effects of ratio and other features of classroom structure appears in Meyer, W., "Staffing Characteristics and Child Outcomes," paper prepared for HEW's Assistant Secretary for Planning and Evaluation under contract No. HEW-100-76-0145. (1977).

3. See Klibanoff, E., "The Enforceability of the FIDCR." Working paper prepared for HEW's Assistant Secretary for Planning and Evaluation. (Cambridge, MA: Abt Associates, 1977).

4. Federal Interagency Day Care Requirements, *op. cit.,* Section III.

5. Federal Interagency Day Care Requirements, *op. cit.,* Section I.B.

6. The following table indicates the number of states with licensing codes which exceed, match or fall short of the FIDCR staff/child ratio provisions:

Number of State* Licensing Codes That:

Age of Children	Exceed FIDCR	Match FIDCR	Fall Short of FIDCR	No Ratio Specified
<2	0	10	28	13
2-3	0	2	44	5
3-4	0	3	46	2
4-5	1	2	46	2
5-6	0	2	47	2
6-10	9	13	25	4

* Includes District of Columbia

Source: U.S. Senate, Senate Finance Committee, *Child Care: Data and Materials, op. cit.*

7. NDCS *Phase II Research Report, op. cit.,* Chapter 2.

Chapter 3: PROCESSES AND OUTCOMES IN DAY CARE

1. Stallings, Wilcox and Travers, *op. cit.*

2. *ibid.*

3. Prescott, E., Summary of Findings, *Assessment of Child-Rearing Environments,* Status Report for OCD Grant No. R219-CG, February, 1973.

4. See, for example, the review in Cronbach, L., *Principles of Educational Psychology.* (New York: Harcourt, Brace Jovanovich, 1977, in press.)

5. Prescott, E. and Milich, C., *School's Out: Group Day Care for the School Age Child.* (Pasadena, California: Pacific Oaks, 1974); Prescott, E., Jones, E., Critchevsky, S., Milich, C. and Haselhoef, E. *Assessment of Child-Rearing Enviornments: An Ecological Approach* (Pasadena, California: Pacific Oaks, 1975).

6. See, for example, Becker, W.C. and Krug, R.S., "A Circumplex Model for Social Behavior in Children," *Child Development,* 1964, Vol. 35, pp. 371-396; Schaefer, E.S., "Converging Conceptual Models for Maternal Behavior and for Child Behavior," in J.C. Glidewell (Ed.), *Parental Attitudes and Child Behavior* (Springfield, IL: Charles C. Thomas, 1961); Baumrind, D. and Black, A.E., "Socialization Practices Associated with Dimensions of Competence in Preschool Boys and Girls," *Child Development,* 1967, Vol. 38, pp. 291-327; Kohn, M.L., and Rosman, B.L., "Relationship of Preschool Social-Emotional Functioning to Later Intellectual Achievement," *Developmental Psychology,* Vol. 6, 1972, pp. 445-452; Kohn, M.L., and Rosman, B.L., "A Social Competence Scale and Symptom Checklist for the Preschool Child: Factor Dimensions, Their Cross-Instrument Generality, and Longitudinal Persistence," *Developmental Psychology,* Vol. 6, 1972, pp. 430-444; Kohn, M.L. and Rosman, B.L., "Cognitive Functioning in Five-Year-Old Boys as Related to Social-Emotional and Background-Demographic Variables," *Developmental Psychology,* Vol. 8, 1973, pp. 227-294; Kohn, M.L., and Rosman, B.L., "Therapeutic Intervention with Disturbed Children in Day Care: Implications of the Deprivation Hypothesis," *Child Care Quarterly,* Vol. 1, 1971, pp. 21-46; Peterson, D.R., "Behavior Problems in Middle Childhood," *Journal of Consulting Psychology,* Vol. 25, 1961, pp. 205-209.

7. See the review by Maccoby, E.E. and Masters, J.C., "Attachment and Dependency." In P. Mussen (Ed.), *Carmichael's Manual of Child Psychology,* 3rd edition, Vol. 2. (New York: Wiley, 1970).

8. Kagan, J. and Kogan, N., "Individual Variations in Cognitive Processes." In Mussen, *op. cit.;* see especially pp. 1300-1301.

9. See, for example, Piaget, J., *Judgment and Reasoning in the Child* (New York: Harcourt, Brace, 1928); Piaget, J., *The Moral Judgment of the Child* (New York: Harcourt, Brace, 1932); Kohlberg, L. "Moral Development and Identification." In H.W. Stevenson (Ed.), *Child Psychology: 62nd Yearbook of the National Society for the Study of Education* (Chicago: University of Chicago Press, 1963).

10. Parten, M.B. "Social Participation Among Preschool Children," *Journal of Abnormal and Social Psychology,* Vol. 27, 1932, pp. 243-269.

11. See, for example, the review by Feshback,S., "Aggression" in Mussen, *op. cit.*

12. Lay, M. and Meyer, W., *Effects of Early Day Care Experience on Subsequent Observed Program Behaviors,* unpublished paper, Syracuse University, 1973. Note, however that the apparent decline in hostility observed during Phase II may be due to changes in coding procedures between the fall and spring. This result will be re-examined during Phase III.

13. See Maccoby and Masters, *op. cit.*

14. Again the decline may be due to changes in the instrument and will be re-examined in Phase III.

15. Kagan and Kogan, *loc. cit.*

16. The PSI is described in the *Preschool Inventory Revised Edition – 1970: Handbook* (Princeton, NJ: Educational Testing Service, 1970). Data from previous samples appear in Gilbert, L.E. and Shipman, V.C., "Technical Report 13: Preschool Inventory" in V.C. Shipman (Ed.), *Disadvantaged Children and Their First School Experiences: ETS Head Start Longitudinal Study Technical Report Series* (Princeton, NJ: Educational Testing Service, 1972); Walker, D.V.,

Bane, M.J. and Bryk, T., *The Quality of the Head Start Planned Variation Data* (Cambridge, MA: Huron Institute, 1973).

17. McCarthy, D., Manual for *McCarthy's Scales of Children's Abilities.* (New York: Psychological Corporation, 1972).

18. The MFFT has been used extensively by Jerome Kagan and his colleagues. See, for example, Kagan, J. "Impulsive and Reflective Children: Significance of Conceptual Tempo," in J.D. Krumboltz (Ed.), *Learning and the Educational Process.* (Chicago: Rand McNally, 1965).

19. See Bentler, P.M. and McClain, J., "A Multitrait-Multimethod Analysis of Reflection-Impulsivity," *Child Development,* 1976, Vol. 47, pp. 218-226; Egeland, B. and Weinberg, [illegible], "The Matching Familiar Figures Test: A Look at its Psychometric Credibility," *Child Development,* 1976, Vol. 47, pp. 483-491; Block, J., Block, J.H., and Harrington, D.M., "Some Misgivings about the Matching Familiar Figures Test as a Measure of Reflection-Impulsivity," *Developmental Psychology,* 1974, Vol. 10, pp. 611-632; Kagan, J. and Messer, S.B., A Reply to "Some Misgivings about the Matching Familiar Figures Test as a Measure of Reflection-Impulsivity," *Developmental Psychology,* 1975, Vol. 11, pp. 244-248; Block, J., Block, J.H., and Harrington, D.M., "Comment on the Kagan-Messer Reply," *Developmental Psychology,* Vol. 11, 1975, pp. 249-252; Ward, W.C., "Technical Report 11: Matching Familiar Figures Test," in V.C. Shipman (Ed.) *op. cit.*

20. Ward, W.C., "Technical Report 13: Motor Inhibition Test," in V.C. Shipman (Ed.), *op. cit.*

21. High/Scope Educational Research Foundation and Abt Associates Inc., *Home Start Evaluation: Final Report Findings and Implications,* (Ypsilanti, MI and Cambridge, MA: High/Scope Educational Research Foundation and Abt Associates, 1976).

22. Schaefer, E.S. and Aaronson, M.R., "Classroom Behavior Inventory: Preschool to Primary" (Bethesda, MD: National Institute of Mental Health, 1966).

23. Kohn, M.L. and Rosman, B.L., "A Social Competence Scale and Symptom Checklist for the Preschool Child: Factor Dimensions, Their Cross-Instrument Generality, and Longitudinal Persistence," *Developmental Psychology,* Vol. 6, 1972, pp. 430-444.

24. *ibid.*

25. Doll, E.A., *The Measurement of Social Competence: A Manual for the Vineland Social Maturity Scale* (Circle Pines, MN: American Guidance Service, 1953).

Chapter 5: PHASE III DESIGN AND DATA COLLECTION

1. The distinction between *quasi-experiments* and *true experiments* is derived from Campbell, D. and Stanley, J., *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand McNally, 1963). In true experiments, the researcher has full control over the conditions to which subjects are exposed; in particular, he can assign subjects randomly to groups that receive different "treatments" to be compared. In quasi-experiments, the pairing of subjects and treatments is outside the researcher's control. Since random assignment is not possible, treatment groups may not be equivalent, necessitating use of statistical adjustments of the kind discussed in Appendix A.

2. In Table 5-1 and in earlier reports, the Phase III design has been presented as a 3x3x2 factorial, with site, ratio, and PSI as independent variables. In fact, this mode of presentation has been employed for reasons of expository clarity and as a device to aid in assignment of centers to treatment and comparison conditions. For example, blocking on PSI has been incorporated in the design to facilitate selection and assignment of centers in such a manner that child characteristics will vary appreciably within treatment groups and sites, but will be equivalent across treatment groups and sites. Such blocking has no direct analytic significance. High and low ratio and PSI blocks will *not* be used as discrete levels in a factorial ANOVA; rather, data will be analyzed by regression, which will take full account of variations in ratio and PSI within, as well as across, blocks of the design. See Appendix A for further discussion.

3. See Chapter 3, Note 19.

4. The RPIS is described in the NDCS *First Annual Report, op. cit.*, pp. 49-50.

5. Shipman, V.C., McKee, J.D. and Bridgeman, B. "Stability and Change in Family Status, Situational and Process Variables and Their Relationship to Children's Cognitive Performance." In V.C. Shipman (Ed.), *op. cit.*

Appendix A: NDCS MEASUREMENT AND ANALYSIS METHODOLOGY

1. Kent, R.N. and Foster, S.L., "Direct observational procedures: methodological issues in naturalistic settings," in A.R. Ciminero, K.S. Calhoun and H.F. Adams (Eds.), *Handbook of Behavioral Assessment* (New York: Wiley, 1977).

2. Cronbach, L.J., Gleser, G.C., Nanda, H. and Rajaratnam, N., *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles* (New York: Wiley, 1972).

3. Kerlinger, F.N. and Pedhazur, E.J., *Multiple Regression in Behavioral Research* (New York: Holt, Rinehart and Winston, 1973).

4. Cohen, J. and Cohen P., *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Hillsdale, NH: Lawrence Erlbaum Associates, 1975).

5. Searle, S.R., *Linear Models* (New York: Wiley, 1971).

6. Cronbach, L.J., Rogosa, D.R., Floden, R.E. and Price, G.G., "Analysis of covariance in nonrandomized experiments: parameters affecting bias" (Stanford, CA: Evaluation Consortium, Stanford University, 1977).

7. Hanuschek, E.A. and Jackson, J.E., *Statistical Methods for Social Scientists* (New York: Academic Press, 1977).

8. *ibid.*

9. Harris, C.W. (Ed.), *Problems in Measuring Change* (Madison, WI: University of Wisconsin Press, 1963).

10. Campbell, D.T., and Erlebacher, A., "How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful," in J. Hellmuth (Ed.), *Compensatory Education: A National Debate, Vol. III. Disadvantaged Child* (New York: Brunner-Mazel, 1970); Cochran, W.G. and Rubin, D.B., "Controlling bias to observational studies: a review," *Sankhya, The Indian Journal of Statistics*, Series A., Volume 35, 1973, Part 4, pp. 417-446; Cronbach, L.J. and Furby, L., "How we should measure 'change' – or should we?" *Psychological Bulletin*, Vol. 74,

1974, pp. 68-80: Kenny, D., "A quasi-experimental approach to assessing treatment effects in the nonequivalent control group design," *Psychological Bulletin*, Vol. 76, 1975; Lord, F.J., "A paradox in the interpretation of group comparisons," *Psychological Bulletin*, Vol. 68, 1967, pp. 304-305; Tucker, L.R., Damarin, F., and Messick, S., "base-free measure of change," *Psychometrika*, Vol. 31, 1966; Werts, C.E. and Linn, R.L., "A general linear model for studying growth," *Psychologlical Bulletin*, Vol. 73, 1970.

11. Kenny, D., *op. cit.*

12. Cook, T.D., "The potential and limitations of secondary evaluations," in M. Apple, N.J. Luffler and B.M. Subhovial (Eds.), *Educational Evaluation: Analysis and Responsibility* (Berkeley, CA: McCutcheon Press, 1974).

13. Bryk, A.S. and Weisberg, H.I., "Use of the nonequivalent control group design when subjects are growing," *Psychological Bulletin*, Vol. 84, 1977, 950-962.

14. Brennan, R.L., *Some Statistical Problems in the Evaluation of Self-Instructional Programs*, doctoral dissertation, Harvard University. (Ann Arbor, Mich.: University Microfilms, 1970, No. 70-23080.)

15. Originally suggested in Bryk, A.S. and Weisberg, H.I., "Value-added analysis: a dynamic approach to the estimation of treatment effects," *Journal of Educational Statistics*, Vol. I, 1976, pp. 127-156, substantially extended in Strenio, J., Bryk, A.S. and Weisberg, H.I., "An Individual Growth Model Perspective for Evaluating Educational Programs," *Proceedings of the 1977 Annual Meeting of the American Statistical Association* (Washington, D.C.: American Statistical Association, in preparation).

16. Goodrich, R.L., "Analytic issues," in NDCS *Phase II Research Report*, *op. cit.*

17. Warner, H., "Robust statistics: a survey and some prescriptions," *Journal of Educational Statistics*, Vol. 1, 1976, pp. 285-312.

18. Haney, W., *Units of Analysis Issues in the Evaluation of Project Followthrough, or There Must Be Heresy in This Same Place* (Cambridge, MA: Huron Institute, 1974).

19. Cronbach, L.J., *Research on Classrooms and Schools: Formulation of Questions, Design, and Analysis* (Stanford, CA: Evaluation Consortium, Stanford University, 1976).

20. *ibid.*

21. *ibid.*

22. Hannan, M.T., *Aggregation and Disaggregation in Sociology* (Lexington, MA: D.C. Heath and Co., 1971); Goodrich, R.L., "National Day Care Study Phase III Interim Report, No. 3, *Phase III Plan for Analysis of Effects.*" (Cambridge, MA: Abt Associates, 197[illegible]).

23. Bracht, G.H., and Glass, G.V., "The external validity of experiments," *American Educational Research Journal*, Vol. 5, 1968 5, 437-474.

Appendix C: THE INFANT CARE SURVEY

1. Recent research reports include: Hock, E., "Alternative Approaches to Child Rearing and their Effects on the Mother-Infant Relationship," Final Report to the Office of Child Development, Grant No. OCD-

CB-490, March, 1976; Kagan, J., Kearsley, R.B. and Zelazo, P.R., "The Effects of Infant Day Care on Psychological Development," paper presented at the meeting of the American Association for the Advancement of Science, Boston, MA, February 1976; Ramey, C. and Mills, P., "Mother-Infant Interaction Patterns as a Function of Rearing Conditions," paper presented at the meeting of the Society for Research in Child Development, Denver, CO, April 1975. Ongoing research efforts include the New York Infant Day Care Study by Mark Golden, the Chicago Child Care Study by Allison Clarke-Stewart, and the Stanford Day Care Study by Sueann Ambron and Mark Everson.

2. See, for example, Ramey & Mills, *op. cit.*, and numerous studies reviewed in Fowler, W., "How Adult/Child Ratios Influence Infant Development," *Interchange,* Vol. 6, No. 1, 1975, pp. 17-31.

3. See, for example, Provence, S. and Lipton, R.C., *Infants in Institutions,* (New York: International Universities Press, 1962); Yarrow, L.J., "Maternal Deprivation: Toward an Empirical and Conceptual Reevaluation," *Psychological Bulletin,* Vol. 58, 1961, pp. 459-490; and other sources cited in Fowler, *op. cit.*